



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 182280

TO: Phuong Bui
Location: REM/2A15/2C18
Art Unit: 1638
Friday, March 17, 2006
Case Serial Number: 10/634548

From: Barb O'Bryen
Location: Biotech-Chem Library
Remsen 1a69
Phone: 571-272-2518


barbara.obryen@uspto.gov

Search Notes

RUSH

THIS PAGE LEFT BLANK

CC The current sequence is that of the thale cress LTT1 phytoI kinase wild-
CC type DNA of the invention.

SQ Sequence 1091 BP; 252 A; 260 C; 223 G; 356 T; 0 U; 0 Other;

Query Match	Score	DB	Length
100.0%	1091	12	1091

Matches 1091; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 AAAAAGATAATACAAATATCATTTTCCTATCTATTTGACTTGCAGATTCTT 60

Oy	61	TCTTCTTCCTTTCTCCTCCTCCGCAAACTCAGTCCCTCGTGCCATGGACGAACCT	120
Ddb	61	TCTTCTTCCTTTCTTCTCCTCCTCCGCAACTCAGTCCCTCGTGCCATGGACGAACCT	120

QY 121 TACCTCTATCCCGATCAATATCAGTGTGTCGGTCCGGGAACAATCTTTGACGCTC 180

Db 121 TACCTCTATCCCGATCAATATCAGTGTGTCGGTCCGGGAACAATCTTTGACGCTC 180

QY	181	ACCGGTTCTGTTCTCCGCGTTCTTAAATTTCTCCGTGTTCAATGGTTGACCGGAA	240
Db	181	ACCGGTTCTGTTCTCCGCGTTCTTAAATTTCTCCGTGTTCAATGGTTGACCGGAA	240

QY	241	300
TGGGCTCTGCTACTAGTTAGCTGCTGCTCTCTGATCTCTTAGCAGTTGCCGCGA		
Db	241	300
TGGGCTCTGCTACTAGTTAGCTGCTGCTCTCTGATCTCTTAGCAGTTGCCGCGA		

QY 301 ATTGCGTGTGCATGACGTGGAGCCACCGTGGCAGTGGTGGAGCATACGCGCTTG 360

Db 301 ATTCGCTGTGCATGACGTGGAGCCACCGTGGCAGTGGTGGAGCATACGCGCTTG 360

QY	361	361	420
TCTTAAGCTTCGAGAGTCTCA	CCAGGAAAGTCATTCA	CAGAGTTTGACGAAAGC	420
Db	TCTTAAGCTTCGAGAGTCTCA	CCAGGAAAGTCATTCA	CAGAGTTTGACGAAAGC 420

Qy	Db
421 TTGTGATATTA	421 TTGTGATATTA
CTCAGGCTGCT	CTCAGGCTGCT
TTTTCGACTTG	TTTTCGACTTG
CGGTGGCCAT	CGGTGGCCAT
TTTCAGCGGAT	TTTCAGCGGAT
CGA 480	CGA 480
421 TTGTGATATTA	421 TTGTGATATTA
CTCAGGCTGCT	CTCAGGCTGCT
TTTTCGACTTG	TTTTCGACTTG
CGGTGGCCAT	CGGTGGCCAT
TTTCAGCGGAT	TTTCAGCGGAT
CGA 480	CGA 480

Qy	481	CCGAGGCTCGATCTTGTGCTGCTTTTGTGCCCTTAGTGAATGAGCTTAAAGCTTGTATTA	540
Db	481	CCGAGGCTCGATCTTGTGCTGCTTTTGTGCCCTTAGTGAATGAGCTTAAAGCTTGTATTA	540

Qy 541 ACGGACTATCCATTCTCCCAATTGCGTCTAATCAATCCGTCACAGAGAGGGGAG 6000

Dp 541 ACGGCTATCATTTTCCCCCAATTGCGTCTAATCAATCCGTCACAGAGAGGGGAG 6000

Qy 601 CAGAGAGTTCCTTAAGGTCCTTGTCTCAAGTTCAGCTCTCTTTCTCTGCGGTTT 660

Dh 601 CAGAGAGTTCCTTAAGGTCCTTGTCTCAAGTTCAGCTCTCTTTCTCTGCGGTTT 660

Qy 661 TCTTCGAGAGATCTCCTTAATCGTAGATCTGTTAGCAATGATGTGTGGCCATG 720

nb 661 TCTTCTGAGAGAGATCTCCTTAATCGTAGATCTGTTAGCAATGATGTGTGGCCATG 720

Qy	721	GAATAGCTGATATAATGAGAGCTAAGTTGGTCAACTAGATACCTTACCAACCCAGAA	780
Dh	721	CAATTCCTGATATATATGCGACCTTACCTTGGCTCAACTAGATACCTTACCAACCCAGAA	780

781 AGAGTTGGGCGAAGCATCTTCATCTTCGGCTTCATCTCCATCGCATTAC 840
 782 |||||
 783 |||||

841 TTTACTATTACTCAAGCCTTGGGTACTTCATCATGAACCTGGGAAACGACCTTGACAGAG 500

901 TAGCAATGGTCTCAATGATGCGCACGGTAGTCGAGTCTACCATTCACCGCATCAATTAG 960

QY 961 AC6CAATATTTGGTTCCTCTGGCTACTATTTTAGCTGCTTATTTTAACTTTTCGATATTT 1020

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100 101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000 1001 1002 1003 1004 1005 1006 1007 1008 1009 1010 1011 1012 1013 1014 1015 1016 1017 1018 1019 1020 1021 1022 1023 1024 1025 1026 1027 1028 1029 1030 1031 1032 1033 1034 1035 1036 1037 1038 1039 1040 1

1000

Db 1081 CAAATGTTCC 1091

RESULT 2	
AAC36634	
ID	AAC36634 standard; DNA; 1088 BP.

AC	AAC36634;
XX	
DT	17-OCT-2000 (first entry)

DE	Arabidopsis thaliana DNA fragment SEQ ID NO: 14519.
XX	
KW	Hybridisation assay; genetic mapping; gene expression control;

KM promoter, termination sequence; **ss**.
XX
OS *Arabidopsis thaliana*.

PN	EP1033405-A2.
XX	
PD	06-SEP-2000.

PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.

PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.

PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.

PR 10-APR-1999; 99US-0130872.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.

PR 23-APR-1999; 99US-0132002E.
PR 28-APR-1999; 99US-0131449D.
PR 30-APR-1999; 99US-0132048D.
PR 30-APR-1999; 99US-0132407D.

PR
06-MAY-1999;
99US-0132487D.
PR
06-MAY-1999;
99US-0132486D.
PR
06-MAY-1999;
99US-0132485D.
PR
06-MAY-1999;
99US-0132484D.

U/-MAY-1999; 99US-0342863.
PK
PR 11-MAY-1999; 99US-034326P.
PR 14-MAY-1999; 99US-0343218P.
PR
PR 14-MAY-1999; 99US-0343210P.
DD

PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134943P.

PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR MAY 1999; 99US-0135652P.

PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222Z.
PR 01-JUN-1999; 99US-0137222Z.

PR 04-JUN-1999; 99US-0137502P.

PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139753P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144864P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 21-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145224P.
PR 23-JUL-1999; 99US-0145276P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.

PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149328P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149466P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149733P.
PR 20-AUG-1999; 99US-0149929P.
PR 20-AUG-1999; 99US-0149930P.
PR 23-AUG-1999; 99US-0149930P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156566P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 16-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161993P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 96.2%; Score 1049; DB 3; Length 1088;
Best Local Similarity 98.8%; Pred. No. 1,3e-304;
Matches 1076; Conservative 0; Mismatches 10; Indels 3; Gaps 2;

[illegible]

0Y	1081	CAATGTTCC	1091
Db	1078	CAATGTTCC	1088
RESULT 3			
ID	ADX46288	standard; cDNA; 1015 BP.	
AC	ADX46288;		
DT	21-APR-2005	(first entry)	
DE	Plant full length insert polynucleotide seqid 21028.		
KM	plant protectant; plant growth regulant; gene therapy; plant;		
KM	recombinant DNA construct; physical array; plant breeding marker;		
KM	cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;		
KM	extreme osmotic condition; pathogen tolerance; pest tolerance;		
KM	growth rate; cell cycle pathway; disease resistance;		
KM	galactomannan production; lignin production; plant growth regulator;		
KM	yield; plant growth; plant development; seed oil; protein yield;		
KM	protein content; gene; ss.		
XX			
OS	Unidentified.		
XX			
PN	US2004034888-A1.		
XX			
PD	19-FEB-2004.		
XX			
PF	28-APR-2003; 2003US-00425114.		
PR	06-MAY-1999; 99US-00304517.		
PR	05-NOV-2001; 2001US-00985678.		
XX			
PA	(LIUJ/) LIU J.		
PA	(ZHOU/) ZHOU Y.		
PA	(KOVA/) KOVALIC D K.		
PA	(SCRE/) SCREEN S E.		
PA	(TABAS/) TABASKA J E.		
PA	(CAOY/) CAO Y.		
XX			
PI	Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JE, Cao Y;		
DR	WPI; 2004-180133/17.		
XX			
PT	New recombinant DNA construct, useful for improving plant tolerance to		
PT	cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or		
PT	pests, for conferring increased resistance to plant disease, or for		
PT	improving yield.		
XX			
PS	Claim 1; SEQ ID NO 21028; 15pp; English.		
XX			
CC	The invention describes a recombinant DNA construct comprising a		
CC	polynucleotide consisting of a sequence encoding an amino acid sequence		
CC	available in electronic form from the US patent office at		
CC	ftp.secdna.uspto.gov/sequence.html?docID:2004034888. The polynucleotide		
CC	of the invention are also useful in physical arrays of molecules and as		
CC	plant breeding markers. The recombinant DNA construct is useful for		
CC	improving plant tolerance to cold, heat, drought, herbicides, extreme		
CC	osmotic conditions, pathogens or pests, for manipulating growth rate in		
CC	plant cells by modification of the cell cycle pathway, for conferring		
CC	increased resistance to plant disease, for producing galactomannan,		
CC	lignin or plant growth regulators, for increasing the rate of homologous		
CC	recombination in plants, for improving yield by modification of		
CC	photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake		
CC	or by providing improved plant growth and development under at least one		
CC	stress condition or for modifying seed oil or protein yield and/or		
CC	content. This sequence represents a plant full length insert		
CC	polynucleotide that can be used in the recombinant DNA construct of the		
CC	invention.		
XX			
Sequence	1015 BP; 228 A; 245 C; 219 G; 323 T; 0 U; 0 Other;		

Query Match 93.0%; Score 1015; DB 13; Length 1015;
 Best Local Similarity 100.0%; Pred. No. 2,2e-284;
 Matches 1015; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy 74 CTTCTCTCTCTCCAACTCAGTTCCTCCGTCAGCAGCAACTTACCTTATCTCC 133
Db 1 CTTCTCTCTCTCCAACTCAGTTCCTCCGTCAGCAGCAACTTACCTTATCTCC 60
Qy 134 GATCAATCATAGTGTGTGCGGTTCGGGACAACTCTTTCAGCACTCCGCTTCTTC 193
Db 61 GATCAATCATAGTGTGTGCGGTTCGGGACAACTCTTTCAGCACTCCGCTTCTTC 120
Qy 194 TCCGCGCTCTGATTTCTTCTCTGTTTATGTTTGAACCGGAATGGGCTCGCTAC 253
Db 121 TCCGCGCTCTGATTTCTTCTCTGTTTATGTTTGAACCGGAATGGGCTCGCTAC 180
Qy 254 TCAGTTACGTCGTCGCTCTCTGATCTCTTTCAGCACTTCGCAATTCGCTGTTGCA 313
Db 181 TCAGTTACGTCGTCGCTCTCTGATCTCTTTCAGCACTTCGCAATTCGCTGTTGCA 240
Qy 314 TGACGTGAGACCAACCGTGGAGTCTTGTTGAGACATACCGCTTGTCTTAACTTGA 373
Db 241 TGACGTGAGACCAACCGTGGAGTCTTGTTGAGACATACCGCTTGTCTTAACTTGA 300
Qy 374 GAGTCTACCAAGGAAAGCTTCAATCAAGATTTGAGAGAAAGCTTGTGCAATATAT 433
Db 301 GAGTCTACCAAGGAAAGCTTCAATCAAGATTTGAGAGAAAGCTTGTGCAATATAT 360
Qy 434 CTCAGTCTGCTTTTCTGACTTGGTGGCAATCTTTCAGCGGATTCAGCCGAGCTCGATA 493
Db 361 CTCAGTCTGCTTTTCTGACTTGGTGGCAATCTTTCAGCGGATTCAGCCGAGCTCGATA 420
Qy 494 CTTTGTCTGCTTTTGTTCGTTAGTGAATGCTTAAAGCTTGTATTAACGACATATCAT 553
Db 421 CTTTGTCTGCTTTTGTTCGTTAGTGAATGCTTAAAGCTTGTATTAACGACATATCAT 480
Qy 554 TTCCCAAAATCGAATGCTTCAATCAAAATCGTCAAGAAAGGAGAGAAAGGTTGCT 613
Db 481 TTCCCAAAATCGAATGCTTCAATCAAAATCGTCAAGAAAGGAGAGAAAGGTTGCT 540
Qy 614 TAAAGTCTCTTGTCTTCTAGCTCTTCTTCTCTGCGTCTTCTTCTGAGAGA 673
Db 541 TAAAGTCTCTTGTCTTCTAGCTCTTCTTCTCTGCGTCTTCTTCTGAGAGA 600
Qy 674 GTCTCTATCGGATGATCTCTGTTAGCAATGATGTGTGCGATGGAATAGCTGATAT 733
Db 601 GTCTCTATCGGATGATCTCTGTTAGCAATGATGTGTGCGATGGAATAGCTGATAT 660
Qy 734 AATGGGACGTAAAGTTGGGTCACTAAGATTCCTTAACCAAGAAAGATTTGGGACG 793
Db 661 AATGGGACGTAAAGTTGGGTCACTAAGATTCCTTAACCAAGAAAGATTTGGGACG 720
Qy 794 AAGCATCTCAATGTTTCTTCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 853
Db 721 AAGCATCTCAATGTTTCTTCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 780
Qy 854 AAGCTTGGTACCTTTCATGAACTGGGAAACGACTTTCAGAGAGATGAGCAATGTCTC 913
Db 781 AAGCTTGGTACCTTTCATGAACTGGGAAACGACTTTCAGAGAGATGAGCAATGTCTC 840
Qy 914 AATGGTGGCAAGGTAGTGGAGTGGCTTACCATCAAGCATCAATTAAGCAATATTTTC 973
Db 841 AATGGTGGCAAGGTAGTGGAGTGGCTTACCATCAAGCATCAATTAAGCAATATTTTC 900
Qy 974 GGTTCCTCTGCTCTATTTTATGCTTATTTAAGTTTTCGATATTAGATTATTCCTTC 1033
Db 901 GGTTCCTCTGCTCTATTTTATGCTTATTTAAGTTTTCGATATTAGATTATTCCTTC 960
Qy 1034 ATAAACGGAATGTATATACGATTTTAAATGATTCGACCTTACAAATGTT 1088
Db 961 ATAAACGGAATGTATATACGATTTTAAATGATTCGACCTTACAAATGTT 1015

```

RESULT 4
 ADX31786
 ID ADX31786 standard; cDNA; 980 BP.
 XX
 AC ADX31786;
 XX
 DT 21-Apr-2005 (first entry)
 XX
 DE Plant full length insert polynucleotide seqid 14606.
 XX
 KW plant protectant; plant growth regulator; gene therapy; plant;
 KW recombinant DNA construct; physical array; plant breeding marker;
 KW cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;
 KW extreme osmotic condition; pathogen tolerance; pest tolerance;
 KW growth rate; cell cycle pathway; disease resistance;
 KW galactomannan production; lignin production; plant growth regulator;
 KW yield; plant growth; plant development; seed oil; protein yield;
 KW protein content; gene; ss.
 XX
 OS Unidentified.
 XX
 PN US2004034888-A1.
 XX
 PD 19-FEB-2004.
 XX
 PF 28-APR-2003; 2003US-00425114.
 XX
 PR 06-MAY-1999; 99US-00304517.
 XX
 PR 05-NOV-2001; 2001US-00985678.
 XX
 PA (Liu/J.) LIU J.
 PA (Zhou/J.) ZHOU Y.
 PA (Kova/J.) KOVALIC D. K.
 PA (Scre/J.) SCREEN S. E.
 PA (Taba/J.) TABASKA J. E.
 PA (Cao/Y.) CAO Y.
 XX
 PI Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JE, Cao Y;
 XX
 DR WPI; 2004-180133/17.
 XX
 PT New recombinant DNA construct, useful for improving plant tolerance to
 PT cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or
 PT pests, for conferring increased resistance to plant disease, or for
 PT improving yield.
 XX
 PS Claim 1; SEQ ID NO 14606; 15pp; English.
 XX
 CC The invention describes a recombinant DNA construct comprising a
 CC polynucleotide consisting of a sequence encoding an amino acid sequence
 CC available in electronic form from the US patent office at
 CC ftp://seqdata.uspto.gov/sequence.html; DocID:2004034888. The polynucleotide
 CC of the invention are also useful in physical arrays of molecules and as
 CC plant breeding markers. The recombinant DNA construct is useful for
 CC improving plant tolerance to cold, heat, drought, herbicides, extreme
 CC osmotic conditions, pathogens or pests, for manipulating growth rate in
 CC plant cells by modification of the cell cycle pathway, for conferring
 CC increased resistance to plant disease, for producing galactomannan,
 CC lignin or plant growth regulators, for increasing the rate of homologous
 CC recombination in plants, for improving yield by modification of
 CC photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake
 CC or by providing improved plant growth and development under at least one
 CC stress condition or for modifying seed oil or protein yield and/or
 CC content. This sequence represents a plant full length insert
 CC polynucleotide that can be used in the recombinant DNA construct of the
 CC invention.
 CC
 SQ Sequence 980 BP; 223 A; 229 C; 215 G; 313 T; 0 U; 0 Other;

Query Match 89.6%; Score 977; DB 13; Length 980;
 Best Local Similarity 100.0%; Pred. No. 5.8e-283;
 Matches 977; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy 115 CAACCTTACCTCTATCTCCGATCATCATCAGTTGTCGGTCCGGGAAACAACCTCTTGA 174
Db 1 CAACCTTACCTCTATCTCCGATCATCATCAGTTGTCGGTCCGGGAAACAACCTCTTGA 60
Qy 175 CGACTCACCGGCTCTGTTCTCCTGCGCTTGTGATTTCTTCTCTGTTTTCATGCTTGA 234
Db 61 CGACTCACCGGCTCTGTTCTCCTGCGCTTGTGATTTCTTCTCTGTTTTCATGCTTGA 120
Qy 235 CCGGAATGGGCTCTGCTACTGATTAAGTGTGCTGCTGCTCTGATCTCTCAGAGTGG 294
Db 121 CCGGAATGGGCTCTGCTACTGATTAAGTGTGCTGCTGCTCTGATCTCTCAGAGTGG 180
Qy 295 CGAGCAATTCGCTGTGCTGATGACGTCGAGGACACGTCGAGCTGCTGTGAGACATACG 354
Db 181 CGAGCAATTCGCTGTGCTGATGACGTCGAGGACACGTCGAGCTGCTGTGAGACATACG 240
Qy 355 CGCTTGTCTTAAGCTTGGAGAGTCTCACCAAGCGAAACGTCATTCACAGTTTGAACA 414
Db 241 CGCTTGTCTTAAGCTTGGAGAGTCTCACCAAGCGAAACGTCATTCACAGTTTGAACA 300
Qy 415 GAAAGCTTGTGCTATATCTCAGGTCGCTTTCGATCTGCGTGGCCATCTTCAAGCG 474
Db 301 GAAAGCTTGTGCTATATCTCAGGTCGCTTTCGATCTGCGTGGCCATCTTCAAGCG 360
Qy 475 GATGACCGAGGCTCGATCTTGTGCTGCTTGTGCTGCTGATGATGCTTAAAGCTTGG 534
Db 361 GATGACCGAGGCTCGATCTTGTGCTGCTTGTGCTGCTGATGATGCTTAAAGCTTGG 420
Qy 535 TTATTAACGACATTCATCCATTTCCCAAAATTTGATGCTATCAATCCGTCACAGAGAG 594
Db 421 TTATTAACGACATTCATCCATTTCCCAAAATTTGATGCTATCAATCCGTCACAGAGAG 480
Qy 595 GGAAGAGGAGAGAGTGTGCTTAAAGTCTTGTCTGATGCTGAGCTTCTTCTTCTG 654
Db 481 GGAAGAGGAGAGAGTGTGCTTAAAGTCTTGTCTGATGCTGAGCTTCTTCTTCTG 540
Qy 655 CGGTTTTCTTCTGAGAGAGTCTCTATCGGATGATCTCGTTAGCAATGATGTGTG 714
Db 541 CGGTTTTCTTCTGAGAGAGTCTCTATCGGATGATCTCGTTAGCAATGATGTGTG 600
Qy 715 GCGATGCAATGCTGATTAATGGAAGCTTAAGTTGGGTCACTAAGTACCTTCAACC 774
Db 601 GCGATGCAATGCTGATTAATGGAAGCTTAAGTTGGGTCACTAAGTACCTTCAACC 660
Qy 775 CAAGAAAGATTTGGGAGAGAGATCTCATGCTTCTGAGCTTCTTCACTCCATCG 834
Db 661 CAAGAAAGATTTGGGAGAGAGATCTCATGCTTCTGAGCTTCTTCACTCCATCG 720
Qy 835 CATTACTTCTTACTCAAGCCTTGGGTACTTCAATGAACTGGGAAACGACTTGC 894
Db 721 CATTACTTCTTACTCAAGCCTTGGGTACTTCAATGAACTGGGAAACGACTTGC 780
Qy 895 AGAAGATGCAATGCTCAATGCTGCGACCGATGCTGATGCTGCTGCTCAACCGATC 954
Db 781 AGAAGATGCAATGCTCAATGCTGCGACCGATGCTGATGCTGCTGCTCAACCGATC 840
Qy 955 AATTAGAGCAATATTTGGTCTCTGCGTCACTTATTTAGCTGCTTATTAAGTTGG 1014
Db 841 AATTAGAGCAATATTTGGTCTCTGCGTCACTTATTTAGCTGCTTATTAAGTTGG 900
Qy 1015 GATATTAATTAATCCCTCATTAACCGAATGATATATATATTAATTAAGTCCG 1074
Db 901 GATATTAATTAATCCCTCATTAACCGAATGATATATATATTAATTAAGTCCG 960
Qy 1075 ACCTTACCAATGTTCC 1091
Db 961 ACCTTACCAATGTTCC 977

```

RESULT 5
ADXS4389
ID ADXS4389 standard; cDNA; 1062 BP.

```

XX AC ADXS4389;
XX XX
XX 21-APR-2005 (first entry)
XX DE Plant full length insert polynucleotide seqid 29129.
XX DE
XX plant protectant; plant growth regulant; gene therapy; plant;
XX recombinant DNA construct; physical array; plant breeding marker;
XX cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;
XX extreme osmotic condition; pathogen tolerance; pest tolerance;
XX growth rate; cell cycle pathway; disease resistance;
XX galactomannan production; lignin production; plant growth regulator;
XX yield; plant growth; plant development; seed oil; protein yield;
XX protein content; gene; ss.
XX
XX Unidentified.
XX OS
XX US2004034888-A1.
XX PN
XX 19-FEB-2004.
XX PD
XX 28-APR-2003; 2003US-00425114.
XX PR
XX 06-MAY-1999; 99US-00304517.
XX PR
XX 05-NOV-2001; 2001US-00985678.
XX PA
XX (LIU/) LIU J.
XX (ZHOU/) ZHOU Y.
XX (KOVA/) KOVALIC D K.
XX (SCRE/) SCREEN S E.
XX (TABAS/) TABASKA J E.
XX (CAO/) CAO Y.
XX PI
XX Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JE, Cao Y;
XX DR
XX WPI; 2004-180133/17.
XX XX
XX New recombinant DNA construct, useful for improving plant tolerance to
XX cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or
XX pests, for conferring increased resistance to plant disease, or for
XX PT improving yield.
XX PT
XX Claim 1; SEQ ID NO 29129; 15bp; English.
XX XX
XX The invention describes a recombinant DNA construct comprising a
XX polynucleotide consisting of a sequence encoding an amino acid sequence
XX available in electronic form from the US patent office at
XX ftp.segdata.uspto.gov/sequence.html?DocId:2004034888. The polynucleotide
XX of the invention are also useful in physical arrays of molecules and as
XX plant breeding markers. The recombinant DNA construct is useful for
XX improving plant tolerance to cold, heat, drought, herbicides, extreme
XX osmotic conditions, pathogens or pests, for manipulating growth rate in
XX plant cells by modification of the cell cycle pathway, for conferring
XX increased resistance to plant disease, for producing galactomannan,
XX lignin or plant growth regulators, for increasing the rate of homologous
XX recombination in plants, for improving yield by modification of
XX photoynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake
XX or by providing improved plant growth and development under at least one
XX stress condition or for modifying seed oil or protein yield and/or
XX content. This sequence represents a plant full length insert
XX polynucleotide that can be used in the recombinant DNA construct of the
XX invention.
XX XX
XX Sequence 1062 BP; 235 A; 264 C; 243 G; 320 T; 0 U; 0 Other;
XX
XX Query Match 62.2%; Score 678.6; DB 13; Length 1062;
XX Best Local Similarity 82.6%; Pred. No. 3.8e-193;
XX Matches 777; Conservative 0; Mismatches 164; Indels 0; Gaps 0;

```

Qy 108 ATGGCAGCAACCTTACCTCTATCTCCGATCATCATCAGTTGTCGGTCCGGGAAACAAC 167
Db 46 ATGGCAGCAACCTTACCTCTATCTCCGATCATCATCAGTTGTCGGTCCGGGAAACAACG 105

KW	protein identification; signal transduction pathway; metabolic; pathway
KV	promoter; termination sequence; corn; ss.
XX	
OS	Zea mays subsp. mays.
XX	
PN	EP1033405-A2.
XX	
PD	06-SEP-2000.
XX	
PF	25-FEB-2000; 2000EP-00301439.
XX	
ER	25-FEB-1999; 99US-0121825P.
ER	05-MAR-1999; 99US-0123180P.
ER	09-MAR-1999; 99US-0123548P.
PR	23-MAR-1999; 99US-0125788P.
PR	25-MAR-1999; 99US-0126264P.
PR	29-MAR-1999; 99US-0126785P.
PR	01-APR-1999; 99US-0127452P.
PR	06-APR-1999; 99US-0128234P.
PR	08-APR-1999; 99US-0128714P.
PR	16-APR-1999; 99US-0129845P.
PR	19-APR-1999; 99US-0130077P.
PR	21-APR-1999; 99US-0130449P.
PR	23-APR-1999; 99US-0130510P.
ER	23-APR-1999; 99US-0130891P.
PR	28-APR-1999; 99US-0131449P.
PR	30-APR-1999; 99US-0132048P.
PR	30-APR-1999; 99US-0132407P.
PR	04-MAY-1999; 99US-0132484P.
PR	05-MAY-1999; 99US-0132485P.
PR	06-MAY-1999; 99US-0132486P.
PR	06-MAY-1999; 99US-0132487P.
PR	07-MAY-1999; 99US-0132863P.
PR	11-MAY-1999; 99US-0134256P.
PR	14-MAY-1999; 99US-0134218P.
PR	14-MAY-1999; 99US-0134219P.
PR	14-MAY-1999; 99US-0134221P.
ER	14-MAY-1999; 99US-0134370P.
PR	18-MAY-1999; 99US-0134768P.
PR	19-MAY-1999; 99US-0134941P.
PR	20-MAY-1999; 99US-0135124P.
PR	21-MAY-1999; 99US-0135353P.
PR	24-MAY-1999; 99US-0135629P.
PR	25-MAY-1999; 99US-0136021P.
ER	27-MAY-1999; 99US-0136352P.
ER	28-MAY-1999; 99US-0136782P.
PR	01-JUN-1999; 99US-0137222P.
PR	03-JUN-1999; 99US-0137528P.
PR	04-JUN-1999; 99US-0137502P.
PR	07-JUN-1999; 99US-0137724P.
ER	08-JUN-1999; 99US-0138094P.
ER	10-JUN-1999; 99US-0138540P.
PR	14-JUN-1999; 99US-0138847P.
PR	14-JUN-1999; 99US-0139119P.
PR	16-JUN-1999; 99US-0139452P.
PR	16-JUN-1999; 99US-0139453P.
PR	17-JUN-1999; 99US-0139492P.
ER	18-JUN-1999; 99US-0139454P.
ER	18-JUN-1999; 99US-0139455P.
PR	18-JUN-1999; 99US-0139456P.
PR	18-JUN-1999; 99US-0139457P.
PR	18-JUN-1999; 99US-0139458P.
PR	18-JUN-1999; 99US-0139459P.
PR	18-JUN-1999; 99US-0139460P.
ER	18-JUN-1999; 99US-0139461P.
PR	18-JUN-1999; 99US-0139462P.
PR	18-JUN-1999; 99US-0139463P.
PR	18-JUN-1999; 99US-0139750P.
PR	21-JUN-1999; 99US-0139763P.
PR	21-JUN-1999; 99US-0139817P.
PR	22-JUN-1999; 99US-0139899P.
PR	23-JUN-1999; 99US-0140033P.
PR	23-JUN-1999; 99US-01400354P.

Db 61 TCTTCTTCTTCTTCTCCTCCTCCAAACTCAGTTCCTCCGTCATGGCAGCAACT 120

CC available in electronic form and data used to

Query Match	Similarity	27.0%	Score 294.6	DB 13	Length 1033
Best Local	Similarity	62.9%	Pred. No. 1.3e-77		
Match	456	Conservative	0	Mismatches 269	Indels 0
				Gaps	0
Qy	294	GCGACGAATTCGCTTTGGCATGACGTCGGAGGACACCGTGGAGCTGCTTGAGACATAC	353		
Db	192	GCGAGCGCCCGGGGGCTGGGGACCTGGCGCGCGACGAGCTGCTCATCACCGCGGGCGGTAC	251		
Qy	354	GCGCTTGCTTAACTTGGAGAGTCTCACCAAGCGAAAGCTATTCAACAGATTGAGC	413		
Db	252	TCCCTGGGCGCGCTTCGACCGGGCTCACCGGCGCGCGCTCATGAAACAGAACCTGAGC	311		
Qy	414	AGAAAGCTTGACATATACCTCAGAGTCTGCTTTTGTATCTGCGGCGCAATCTTTCAGC	473		
Db	312	AGAAATAATTGACATGCTCTTCGGAGAGTCTGCTTATGATCTTCTCGGCCATTAATCACT	371		
Qy	474	GGATGACCGAGGCTCGATCTTGTCTGCTTTTGTTCGTTAGTAAATGGCTTAAAGCTT	533		
Db	372	AATTGACAGAAAGCAAGGTTCTTTCGGCAATGTCGGTTGCTGAATGCAATTAAGGCTT	431		
Qy	534	GTTATTAAACGACTATCCATTTTCCCAATTGATGCTATATCAATCCGTCAAGAGAA	593		
Db	432	CTGACCTTACGGGCTCCGCTTTCACATGATGAAGCTATGTAATAATCGGTGACCGGTGA	491		
Qy	594	GGGAGAGAGAGAGAGTGGCTTAAAGTCCCTTGTCTACGTTCTAGCTCTTCTTCTCT	653		
Db	492	GGAAACAGAGAAATGCTGAGAGGTCTCTGTATGATGATGCTGCTGCTGAGT	551		
Qy	654	GCGGTTTTCTTCTGAGAGAGTCTCTATCGGTATGATCTCGTTAGCAATGATGTGGT	713		
Db	552	GTTTGGTCTTCTGGCGCAGCTCTCTATGGGATGTTGTTGTTGCTGATGATGATGGT	611		
Qy	714	GGCCTGGAATGCTGATATATATGGGAGTAAAGTTGGGTCACTTAAGTAACTTACAC	773		
Db	612	GGTGATGCTTGTGCTGACATGTTGGAGAAAGTATGGCTCGCAGAGCTGCATTTCAAT	671		
Qy	774	CCAAGAAAGAGTGGGAGAGAGATCTTCATGTTCACTTGGGCTTCTCATCTCCATC	833		
Db	672	GAGAACAGAGCTGATAGAGAGATCTTCATATGTTCTTCTGCTTCTTCTTATCTGCA	731		
Qy	834	GCATTACTTTACTATTAATCTCAAGCCTTGGGATCTTCAATGAACTGGGAAAGCACTTG	893		
Db	732	CTGATGCTTGTCTACTTCTTCTGCTTGGTACTTCACTGTCTGCTGGGATCTGGACAT	791		
Qy	894	CAGAGAGTACGATGCTCTCAATGTCGCCAAGGATGTCAGTCCGTACCCATCACCAT	953		
Db	792	GGTAAACCTGGCTCTTGTGCAATTAACAGCTATCTGATGAGATGTATTTCTGTCAATAT	851		
Qy	954	CAATTAGACGAATATTTGCGTCTCTGCTACATATTTAGCTTATTAAGTTTC	1011		
Db	852	GTTGAGATGACATATCTCGTCTTCTTGGGCCACATGTTGGCAAGCTCATCTGTATTT	911		
Qy	1014	GGATTA 1018			
Db	912	GGCTTA 916			
RESULT 10					
ADX47586					
ID ADX47586 standard; cDNA, 1085 BP.					
XX ADX47586;					
DT 21-APR-2005 (first entry)					
XX Plant full length insert polynucleotide seqid 22326.					
XX plant protease; plant growth regulant; gene therapy; plant;					
XX recombinant DNA construct; physical array; plant breeding marker;					
XX cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;					
XX extreme osmotic condition; pathogen tolerance; pest tolerance;					
XX growth rate; cell cycle pathway; disease resistance;					

KW	galactomannan production; lignin production; plant growth regulator;
KM	yield; plant growth; plant development; seed oil; protein yield;
KN	protein content; gene; ss.
XX	Unidentified.
OS	
XX	US2004034888-A1.
PN	
XX	
PD	19-FEB-2004.
XX	
XX	28-APR-2003; 2003US-00425114.
PF	
XX	06-MAY-1999; 99US-00304517.
PR	05-NOV-2001; 2001US-00985678.
XX	
PA	(LIU/J) LIU J.
PA	(ZHOU/) ZHOU Y.
PA	(KOVA/) KOVALIC D K.
PA	(SCRE/) SCREEN S E.
PA	(TABA/) TABASKA J E.
PA	(CAO/) CAO Y.
XX	
XX	
PI	Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaeka JE, Cao Y;
DR	WPI; 2004-180133/17.
XX	
PT	New recombinant DNA construct, useful for improving plant tolerance to
PT	cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or
PT	pests, for conferring increased resistance to plant disease, or for
PT	improving yield.
XX	
PS	Claim 1; SEQ ID NO 22326; 15pp; English.
XX	
CC	The invention describes a recombinant DNA construct comprising a
CC	polynucleotide consisting of a sequence encoding an amino acid sequence
CC	available in electronic form from the US patent office at
CC	ftp:seeddata.uspto.gov/sequence.html?docID:2004034888. The polynucleotide
CC	of the invention are also useful in physical arrays of molecules and as
CC	plant breeding markers. The recombinant DNA construct is useful for
CC	improving plant tolerance to cold, heat, drought, herbicides, extreme
CC	osmotic conditions, pathogens or pests, for manipulating growth rate in
CC	plant cells by modification of the cell cycle pathway, for conferring
CC	increased resistance to plant disease, for producing galactomannan,
CC	lignin or plant growth regulators, for increasing the rate of homologous
CC	recombination in plants, for improving yield by modification of
CC	photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake
CC	or by providing improved plant growth and development under at least one
CC	stress condition or for modifying seed oil or protein yield and/or
CC	content. This sequence represents a plant full length insert
CC	polynucleotide that can be used in the recombinant DNA construct of the
CC	invention.
XX	
Sequence	1085 BP; 178 A; 307 C; 298 G; 302 T; 0 U; 0 Other;
Query Match	27.0%; Score 294.6; DB 13; Length 1085;
Best Local Similarity	62.9%; Pred. No. 1.4e-77;
Matches	456; Conservative 0; Mismatches 269; Indels 0; Gaps 0
QY	294 GCGAGCAATTCGCTTGCATGACGTGGAGACCAACCGTGGACAGTCTTGGTGGAGCATAC 353
DB	260 GCGACGCCCGGGCGCTGGCGGACACTGCGCGCCACAGCTGCTCATACCGCGGGCGCTAC 319
QY	354 GCGCTTGTCTTAAGCTTCGAGAGTCTCCACGAAGCAAAAGTCATTCACAGATTGAGC 413
DB	320 TCCTGTGTGGCGCGCTTCGACGGGCTACCGGCGCGCGGCTCATCGAACAACCTGAGC 379
QY	414 AGAAAGCTTGTGCATATACCTCTCAGAGTCTGCTTTTGTGATTGCGTGGCCAACTTCACG 473
DB	380 AGAAATAATTTGCAATGTCCTCTCGAGATCTCTGTTACATGCTCTCTCGGCACATATTCAGT 439
QY	474 GGAATGACCGGAGCTCATATTCTTGTCTTGTTCGTTAGTAAAGCTTAAGGCTT 533
DB	440 AATTCGACAGAGCAACGGTGTCTTGGCGGCAATGTCCTGTCGAACTGCATATTAAGGCTT 499

Qy	534	GTATATACGGACTATTCATCTTCCCAATTCAGATGATTAATTCGTCAAGAA	593
Db	500	CTGACTTACGGGCTCCGCTTTCACATGATGAGCTATGATAAATCCGTACCCGTGAA	559
Qy	594	GGGAGAGCAGAAAGTGTCTTAAAGTCTCTTGTCTCACTGTTCTACGCTCTTCTTCTCT	653
Db	560	GGAAAACAGAGAAATTCGTAGAGTCTCTGTACTATGTCATCTTGTCTGTCTGTCACT	619
Qy	654	GCGGTTTTCTTCTGGAGAGTCTCTATTCGGTATGATCTCGTTAGCAATGATGTGTGT	713
Db	620	GTTTTGTCTTCTGTGGCGCAAGTCTCTATTTGGAGATTTGTTGTGTGCATGATGATGTGT	679
Qy	714	GGCGATGGAATAGCTGATATTAATGGGACGTAACTTTGGGTCACTAAGATACCTTACAC	773
Db	680	GGTGTGTGCTTTGTCTGACATGTTGTGGAGAAAGTATGTGCTCGCAACCTGCAATTCAT	739
Qy	774	CCAGAAAGATTTGGGCAAGAAAGATCTCAATGTTCACTTCGTGTTCTTCACTCCATC	833
Db	740	GAGAACAAAGCTGGATGAGAAAGATCTCAATGTTCAATTTCTGGCTTCTTCTATCTGCA	799
Qy	834	GCATTACTTTACTATTACTCAAGCTTGGGTACTTCAATGAATCTGGGAAACGACTTGG	893
Db	800	CTGATCTGTTTCTACTTCTCTTGGCTTGGTACTTCACTGTCGTGCGGATCTGGCACTT	859
Qy	894	CAGAGATGACAAATGTCTCAATGTTGCCACGATGATGAGTGCCTAACCCATCCAGAT	953
Db	860	GGTAAACTGGGCTCTTGTTCATTTACAGTACTGATGAGATGATTTCTGTCAATGAT	919
Qy	954	CAATTAGACGACAAATTTGGTTCCTCTGGTACTATTTTAACTGCTTATTTAAATTTC	1013
Db	920	GTTTGATGACAAATATCTCGTCTCTTTGGCCACCAATGTTGGCAGCTATCTGTATTTT	979
Qy	1014	GGATA 1018	
Db	980	GGCTA 984	

RESULT 11
ADX48515
ID ADX48515 standard; cDNA; 798 BP.

AC	ADX48515;
XX	
DT	21-APR-2005 (first entry)

DE Plant full length insert polynucleotide seqid 23255.

KM plant protectant; plant growth regulator; gene therapy; plant;
KM recombinant DNA construct; physical array; plant breeding maker;
KM cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;
KM extreme osmotic condition; pathogen tolerance; pest tolerance;
KM growth rate; cell cycle pathway; disease resistance;
KM galactomannan production; lignin production; plant growth regulator;
KM yield; plant growth; plant development; seed oil; protein yield;
KM protein content; gene; ss.

OS Unidentified.

PN US2004034888-A1.

PD 19-FEB-2004

28-APR-2003; 2003US-00425114.

06-MAY-1999; 99US-00304517.

05-NOV-2001; 2001US-00985678.

PA (LIU/) LIU J.
PA (ZHOU/) ZHOU Y.
PA (KOVA/) KOVALIC D K.
PA (SCRE/) SCREEN S E.
PA (TABA/) TABASKA J E.

PA (CAOY/) CAO Y.
XX
XX
PI Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JB, Cao Y,
XX
DR WPI, 2004-180133/17.

PT New recombinant DNA construct, useful for improving plant tolerance to PT cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or PT pests, for conferring increased resistance to plant disease, or for PT improving yield.

PS Claim 1; SEQ ID NO 23255; 15pp; English.

The invention describes a recombinant DNA construct comprising a polynucleotide consisting of a sequence encoding an amino acid sequence available in electronic form from the US patent office at <http://seqdata.uspto.gov/sequence.html?DocID:2004034886>. The polynucleotide of the invention are also useful in physical arrays of molecules and as plant breeding markers. The recombinant DNA construct is useful for improving plant tolerance to cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or pests, for manipulating growth rate in plant cells by modification of the cell cycle pathway, for conferring increased resistance to plant disease, for producing galactosemannan, laminin or plant growth regulators, for increasing the rate of homologous recombination in plants, for improving yield by modification of photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake or by providing improved plant growth and development under at least one stress condition or for modifying seed oil or protein yield and/or content. This sequence represents a plant full length insert polynucleotide that can be used in the recombinant DNA construct of the invention.

Sequence 798 BP; 159 A; 195 C; 223 G; 221 T; 0 U; 0 Other;

Query Match 26.5%; Score 289; DB 13; Length 798;

Matches 451; Conservative 0; Mismatches 270

Matches 451; Conservative 0; Mismatches 270; Indels 0; Gaps 0.

Qy	286	CAGCAGTTGCCAGCAATTCCGTGTGGCANTGAGTGGAGCCACCCTGGGAGCTGTGGT	345
Db	2	CGGCGCGCGCGCGCGCAGCGGTGTGCGAGAGCGAGCGGTCACTGTGCTCATACCGCCG	61
Qy	346	GAGCATACGCGCTTGTCTTAAAGCTTCGAGAGTCTCAACCAAGCGAAAGTCAATTCACAG	405
Db	62	GGGCTTACTCTTGTGGCGGTCTTCGACGAGCTCACCGAGCGCGCGCTCATCGAAAGA	121
Qy	406	GTTTGAAGCAAAAGCTTGTGCAATAACTCAGGTCTGCTTTTGTGATCTTGCTGGCCAA	465
Db	122	GCTTGAAGCAGAAAGGTGTGCAAGTCAATCCGCGCTCTGTTCATGTCACTTTGGCCCC	181
Qy	466	TCTTGAGCGGATTCGACCCGAGGCTGCATACTTTGTCTCTTTTGTCCGTTAGTATGGCT	525
Db	182	TGTTGAGCAATTCGACAGAGACAGGTATTTGGCGCGGTGTGTCGGTTCCTAACTCCA	241
Qy	526	TAAAGCTGTATTAACGAGCTATCCATTTCCCGAAATTCAGATCTAATATCCGTCA	585
Db	242	TGAGGCTTCTGATATATATGACTCCGCTCTCTCACTGATAAAGCTCTGGTAAATTCAGTGA	301
Qy	586	CAGAAGAGGAGAGAGAGAGAGTGTCTTAAAGGTCTTGTGTCTAGTCTTACGCTTTC	645
Db	302	CACGTGAGGAAACCAAGAGAAATTCCTGAGAGTCCACTATATATGTCTGTGGCTGC	361
Qy	646	TTTTCTCTGCGGTTTTCTTCTGAGAGAGTCTTCTATTCGGTATGATCTCGTTAGCAATGA	705
Db	362	TGTTGAGCCTTTTAACTCTTCTGCGGAGATCCGCCCATCGGGAATCGTCTCTGTTCGATGA	421
Qy	706	TGTGTGTGGCGATGGAATAGCTGATATATATGAGGAGTAAAGTTTGGGTCACTAAGATAC	765
Db	422	TAAAGCGGTGGCAATGTTTTTGTCTGACATGTTTGGAGAGGATATGGCTACGCAAGCTGC	481
Qy	766	CTTACAAACCAAGAAAGATTGGCGAGAGACATCTCATGTTCACTTTCGCTTTTCA	825
Db	482	CATTCAATCGAAGAAAGACTGGGCGGGAGACATCGATGTTCAATTTCTGATTTCTCTGC	541

QY 826 TCTCATGCACTTACTTACTTACTTCAAGCCTTGGTACCTTCACATGAACTGGGAAA 885
 DB 542 TGTCCGCAATGATGATGCTTACTTCTCAAGCCTGGGTACATGATGATTTCTGGGAG 601
 QY 886 CGACCTTCAGAGAGATGACCAATGCTCTCAATGCTGCCAGAGTACGCTGACCA 945
 DB 602 AGGACCTGTTAGTGGCGCTTGTGACATGACGACAGTATGAGTGGCTTCTG 661
 QY 946 TCACCGATCAATTAGACCAATATTTGCTTCTGCTGCTACTATTTTACGCTTATT 1005
 DB 662 TGACCGAAGTTGTAGATGACAAATATCTGTTCTTGGCCACCATGCTGAGCTTTTC 721
 QY 1006 T 1006
 DB 722 T 722
 RESULT 12
 ADX51811
 ID ADX51811 standard; cDNA, 976 BP.
 XX ADX51811;
 AC
 XX 21-APR-2005 (first entry)
 DT
 XX
 DB Plant full length insert polynucleotide seqid 26551.
 XX
 XX plant protectant; plant growth regulant; gene therapy; plant;
 KW recombinant DNA construct; physical array; plant breeding marker;
 KW cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;
 KW extreme osmotic condition; pathogen tolerance; pest tolerance;
 KW growth rate; cell cycle pathway; disease resistance;
 KW galactomanan production; lignin production; plant growth regulator;
 KW yield; plant growth; plant development; seed oil; protein yield;
 KW protein content; gene; ss.
 XX
 OS Unidentified.
 XX
 XX US2004034888-A1.
 PN
 XX
 PD 19-FEB-2004.
 XX
 XX 28-APR-2003; 2003US-00425114.
 PE
 XX 06-MAY-1999; 99US-00304517.
 PR
 XX 05-NOV-2001; 2001US-00985678.
 XX
 XX (LIUJ/) LIU J.
 PA (ZHOU/) ZHOU Y.
 PA (KOVA/) KOVALIC D K.
 PA (SCRE/) SCREEN S E.
 PA (TABAA/) TABASKA J E.
 PA (CAOY/) CAO Y.
 XX
 XX Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JE, Cao Y;
 PI
 DR WPI: 2004-180133/17.
 XX
 XX New recombinant DNA construct, useful for improving plant tolerance to
 PT cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or
 PT pests, for conferring increased resistance to plant disease, or for
 PT improving yield.
 XX
 XX Claim 1; SEQ ID NO 26551; 15pp; English.
 PS
 CC The invention describes a recombinant DNA construct comprising a
 CC polynucleotide consisting of a sequence encoding an amino acid sequence
 CC available in electronic form from the US patent office at
 CC ftp.segdata.uspto.gov/sequence.html?docID:2004034888. The polynucleotide
 CC of the invention are also useful in physical arrays of molecules and as
 CC plant breeding markers. The recombinant DNA construct is useful for
 CC improving plant tolerance to cold, heat, drought, herbicides, extreme

CC osmotic conditions, pathogens or pests, for manipulating growth rate in
 CC plant cells by modification of the cell cycle pathway, for conferring
 CC increased resistance to plant disease, for producing galactomanan,
 CC lignin or plant growth regulators, for increasing the rate of homologous
 CC recombination in plants, for improving yield by modification of
 CC photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake
 CC or by providing improved plant growth and development under at least one
 CC stress condition or for modifying seed oil or protein yield and/or
 CC content. This sequence represents a plant full length insert
 CC polynucleotide that can be used in the recombinant DNA construct of the
 CC invention.
 CC
 XX
 SQ Sequence 976 BP; 149 A; 300 C; 280 G; 247 T; 0 U; 0 Other;
 Query Match 25.6%; Score 278.8; DB 13; Length 976;
 Best Local Similarity 61.6%; Pred. No. 7.4e-73;
 Matches 445; Conservative 0; Mismatches 277; Indels 0; Gaps 0;
 QY 294 GCGAAGATTGCTGTTGATGACGTCGAGCCACCGTGGCAGTCTTGTGAGCATAC 353
 DB 204 GCGGCTTCATGCTGCTGCGGAGCGCGGCCACGCTGCTGCACCGCGGGGCATAC 263
 QY 354 GCGCTTGTCTTAACTTGAAGTCTCAACGAGGAAAGCTTCAACAGAGTTGAGC 413
 DB 264 TCCCTCGTGGCGGCTTCAACGCGCTCAACGAGCGCGCTCGTCAACAGAGCTTGAAGC 323
 QY 414 AGAAAGCTTGTGATTAATCTCAAGTCTGTTTGTGATCTTGGTGGCCAAATTTGAGC 473
 DB 324 AGGAAGTTGTGATTAATCTCAAGTCTGTTTGTGATCTTGGTGGCCAAATTTGAGC 383
 QY 474 GGATGACCGAGGCTTCATTAATCTTGTGCTTGTGCTTGTGATTAATCTTGTGAGCTT 533
 DB 384 AATTCTACAGTGAACGCTTCTTGTGACGAGTATGCTTGTGCTTGTGATTAATCTTGTGAGCTT 443
 QY 534 GTTATTAACGAGTATCAATCTTCCCAATTCATGATTAATCAATTCGTCACAGAGAA 593
 DB 444 CTCACCTACGAGGCTTGTGATTAATCTTCCCAATTCATGATTAATCAATTCGTCACAGAGAA 503
 QY 594 GGGAGAGCAAGAGTGTCTTAAGTCTTGTGCTTGTGCTTGTGATTAATCTTGTGAGCTT 653
 DB 504 GGAAGAGCAAGAGTGTCTTAAGTCTTGTGCTTGTGCTTGTGATTAATCTTGTGAGCTT 563
 QY 654 GCGGTTTCTTCTGAGAGAGTCTCTTAAGTCTTGTGCTTGTGCTTGTGATTAATCTTGTGAGCTT 713
 DB 564 GTTCTAGTCTTCTTGTGAGAGTCTCTTAAGTCTTGTGCTTGTGCTTGTGATTAATCTTGTGAGCTT 623
 QY 714 GCGCATGGAATAGTGTATTAATGAGAGTATGAGTGTGCTTGTGCTTGTGATTAATCTTGTGAGCTT 773
 DB 624 GTTACGAGCTTGTGATTAATGAGAGTATGAGTGTGCTTGTGCTTGTGATTAATCTTGTGAGCTT 683
 QY 774 CCAAGAAAGAGTGTGAGAGAGTCTCAATGAGTGTGCTTGTGCTTGTGATTAATCTTGTGAGCTT 833
 DB 684 AAGAAAGAGTGTGAGAGAGTCTCAATGAGTGTGCTTGTGCTTGTGATTAATCTTGTGAGCTT 743
 QY 834 GCATTACTTCTTACTTACTTCAAGCTTGGTACCTTCAATGAGTGTGCTTGTGAGCTT 893
 DB 744 CTGATGCTTCTTACTTACTTCAAGCTTGGTACCTTCAATGAGTGTGCTTGTGAGCTT 803
 QY 894 CAGAGAGTGAATGATGCTCAATGAGTGTGCTTCAATGAGTGTGCTTGTGAGCTT 953
 DB 804 GGTAAAGTGTGCTTCAATGAGTGTGCTTCAATGAGTGTGCTTGTGAGCTT 863
 QY 954 CAATTAGACGACATATTTGCTTGTGCTTCAATGAGTGTGCTTGTGAGCTT 1013
 DB 864 GTTGTAGATGACATATTTGCTTGTGCTTCAATGAGTGTGCTTGTGAGCTT 923
 QY 1014 GG 1015
 DB 924 GG 925
 RESULT 13
 ADX51714

XX	ID	ADXS1714	standard; cDNA; 937 BP.
XX	AC	ADXS1714;	
XX	DT	21-APR-2005	(first entry)
XX	DE	Plant full length insert polynucleotide seqid 26454.	
XX	KM	plant protectant; plant growth regulant; gene therapy; plant;	
XX	KM	recombinant DNA construct; physical array; plant breeding marker;	
XX	KM	cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;	
XX	KM	extreme osmotic condition; pathogen tolerance; pest tolerance;	
XX	KM	growth rate; cell cycle pathway; disease resistance;	
XX	KM	galactomannan production; lignin production; plant growth regulator;	
XX	KM	yield; plant growth; plant development; seed oil; protein yield;	
XX	KM	protein content; gene; ss.	
XX	OS	unidentified.	
XX	PN	US2004034888-A1.	
XX	PD	19-FEB-2004.	
XX	PF	28-APR-2003; 2003US-00425114.	
XX	PR	06-MAY-1999; 99US-00304517.	
XX	PR	05-NOV-2001; 2001US-00985678.	
XX	PA	(LIUJ/) LIU J.	
XX	PA	(ZHOU/) ZHOU Y.	
XX	PA	(KOVA/) KOVALIC D K.	
XX	PA	(SCRE/) SCREEN S E.	
XX	PA	(TABAA/) TABASKA J E.	
XX	PA	(CAOY/) CAO Y.	
XX	PI	Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JE, Cao Y;	
XX	DR	WPI; 2004-180133/17.	
XX	PT	New recombinant DNA construct, useful for improving plant tolerance to	
XX	PT	cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or	
XX	PT	pests, for conferring increased resistance to plant disease, or for	
XX	PT	improving yield.	
XX	PS	Claim 1; SEQ ID NO 26454; 15pp; English.	
XX	CC	The invention describes a recombinant DNA construct comprising a	
XX	CC	polynucleotide consisting of a sequence encoding an amino acid sequence	
XX	CC	available in electronic form from the US patent office at	
XX	CC	ftp.begdata.uspto.gov/sequence.html?DocID:2004034888. The polynucleotide	
XX	CC	of the invention are also useful in physical arrays of molecules and as	
XX	CC	plant breeding markers. The recombinant DNA construct is useful for	
XX	CC	improving plant tolerance to cold, heat, drought, herbicides, extreme	
XX	CC	osmotic conditions, pathogens or pests, for manipulating growth rate in	
XX	CC	plant cells by modification of the cell cycle pathway, for conferring	
XX	CC	increased resistance to plant disease, for producing galactomannan,	
XX	CC	lignin or plant growth regulators, for increasing the rate of homologous	
XX	CC	recombination in plants, for improving yield by modification of	
XX	CC	photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake	
XX	CC	or by providing improved plant growth and development under at least one	
XX	CC	stress condition or for modifying seed oil or protein yield and/or	
XX	CC	content. This sequence represents a plant full length insert	
XX	CC	polynucleotide that can be used in the recombinant DNA construct of the	
XX	CC	invention.	
XX	SQ	Sequence 937 BP; 167 A; 261 C; 277 G; 232 T; 0 U; 0 Other;	
XX	Query Match	25.2%; Score 275.4; DB 13; Length 937;	
XX	Beet Local Similarity	62.3%; Pctd No. 7.7e-72;	
XX	Matches 449; Conservative	0; Mismatches 271; Indels 1; Gaps 1	

286 CAGCAGTTGCGACGAATTGCTTTGTCATGACGTGAGCCACCGTGGCAGTGTCTTGATG 345

Db	158	CGGCGCGGCGCGCCGCCACGCGGTGCTGACGAGCGGGGCGGTCACTGTCATCAACCGCG	217
Qy	346	GAGCATACGCGCTTGTCATTAAGCTTCGAGAGTCTCACCAAGCGMAAGTCAATTAACAGA	405
Db	218	GCGCTTACTCTTGTGTGGCGCTCTTCCACGAGCTCAACGACGGCGGCTCATCGAANA	277
Qy	406	GTTTGAGCAGAAAGCTGTGCTATTAATCTTCAGGTCTGCTTTTGTGTAATTGCGGCCAA	465
Db	278	GCTTGAGAGAGAGAGTGTGTGCAAGTGTGTCGCGCGGCTCTGTATGTCATCTTGGCCCC	337
Qy	466	TCTTCAGCGGATCGACCGAGGCTGATTACTTGTGCTTTTGTTCGGTATGAAATGGCT	525
Db	338	TGGTCAGAAATTCACAGAAACAGGATTTGTCGCGGGTGTCCGCTCTGAACCTCCA	397
Qy	526	TAAAGCTGTATTAAACGGAATATCCATTTCCCAAAATTCAGATCTAATCAATCCGCTCA	585
Db	398	TGAGGCTTCTGATATATGAGACTCCGCTCTACATGATGAAGCTGTGTAATAATCAGAGA	457
Qy	586	CAAGAAAGGAGAGCAGAGAGATTGCTTAAAGTCTTTGTCTTACGTTTACGCTCTTC	645
Db	458	CACGTGAAGGAAAAACAGAGGAATTTGCTGAGAGG-CCACTTATTTATGTCCTGTGTCTGC	516
Qy	646	TTTTCTCTGCGGTTTTCTCTCGAGAGAGTCTCCATTCGATATGATCTCGTTAGCAATGA	705
Db	517	TGTTACGCGTTTATGCTCTTCTGCGTGAAGTCCCAATCGGAATGCTCTCTTGTCATGA	576
Qy	706	TGTGTGTGTGCGATGATATAGCTGATATTAATGGAACGTAACTTTGGGTCACTAAGATAC	765
Db	577	TGAGGGGTGGGAGATGGTTTGCTGACATTTGTTGGAAGAGGTATAGGCTCAGCGAAGCTGC	636
Qy	766	CTTACAAACCCAGAAAGATTGGGCGAGAGACATCTCAATTTTCACTCTCGGCTTTCTCA	825
Db	637	CATTCAAATCGAAGAAAGCTGGGCGCGGAGACATCTCGATTTTCAATTTCTGTTTCTGC	696
Qy	826	TCTCATGTGCAATTAATTAATTAATCAAGCTTGGGTACCTTCACTCACTGAATCGGAGAA	885
Db	697	TGTCGCGATATGATGCTCTACTTCTCAACCTCGGGTTAATGATATGTTATTCGGGAAG	756
Qy	886	CGACCTTGACAGAGATGACAAATGCTCTCAATGTGTGCGCACGGTATGTCAGTTCGCTACCA	945
Db	757	AGGCATCTGTGTAGCTGGCGCTTGTGACATGAGACGACAGTATGTGAGATGCGTTCGTG	816
Qy	946	TCACCGATCAATTTAGACGACAATATTTTCGTTCCCTGCGTACTAATTTAGCTGCTTAT	1005
Db	817	TGACCGAAGTTGTATGATGACAAATATCTGTTCTTTGGCCACCAATGCTGTAGCTTTTC	876
Qy	1006	T 1006	
Db	877	T 877	
RESULT 14			
ADXS0931			
ID	ADXS0931	standard; cDNA; 723 BP.	
AC	ADXS0931;		
XX	21-APR-2005	(first entry)	
DT	Plant full length insert polynucleotide seqid 25671.		
XX			
XX	plant proteicant; plant growth regulant; gene therapy; plant;		
KW	recombinant DNA construct; physical array; plant breeding marker;		
KW	cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;		
KW	extreme osmotic condition; pathogen tolerance; pest tolerance;		
KW	growth rate; cell cycle pathway; disease resistance;		
KW	galactomannan production; lignin production; plant growth regulator;		
KW	yield; plant growth; plant development; seed oil; protein yield;		
XX	protein content; gene; ss.		
OS	Unidentified.		
XX			
PN	US2004034888-A1.		

PS Claim 1; SEQ ID NO 26498; 15pp; English.

CC The invention describes a recombinant DNA construct comprising a
 CC polynucleotide consisting of a sequence encoding an amino acid sequence
 CC available in electronic form from the US patent office at
 CC ftp.segdata.uspto.gov/sequence.html?docID:2004034888. The polynucleotide
 CC of the invention are also useful in physical arrays of molecules and as
 CC plant breeding markers. The recombinant DNA construct is useful for
 CC improving plant tolerance to cold, heat, drought, herbicides, extreme
 CC osmotic conditions, pathogens or pests, for manipulating growth rate in
 CC plant cells by modification of the cell cycle pathway, for conferring
 CC increased resistance to plant disease, for producing galactomannan,
 CC lignin or plant growth regulators, for increasing the rate of homologous
 CC recombination in plants, for improving yield by modification of
 CC photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake
 CC or by providing improved plant growth and development under at least one
 CC stress condition or for modifying seed oil or protein yield and/or
 CC content. This sequence represents a plant full length insert
 CC polynucleotide that can be used in the recombinant DNA construct of the
 CC invention.

XX Sequence 753 BP; 154 A; 180 C; 206 G; 213 T; 0 U; 0 Other;

Query Match 25.2%; Score 274.8; DB 13; Length 753;
 Best Local Similarity 63.4%; Pred. No. 1e-71;
 Matches 420; Conservative 0; Mismatches 242; Indels 0; Gaps 0;

QY 345 GGAGCATCGGCGCTTGTCTTAAGCTTCGAGAGCTCACCAAGCCGAAAGTCATTCACAG 404
 DB 16 GCGCGCTTACTCTGTGTGGCGCTTTCGACAGCTCACCGAGCGCGCTCATTCGAAAG 75
 QY 405 AGTTGACGAGAAAGCTTGTGATATACCTCAGGCTGCTTTTGTGTAATTGCGTGCCA 464
 DB 76 AGCTTGACGAGAAAGGTTGTGACGCTGCTCGCGCTCCTGTTGATGATCTTGGCCC 135
 QY 465 ATCTTACGGGATCGACCGAGGCTGATATTGCTGCTTTTGTTCGTTAGTGATGCG 524
 DB 136 CTGTTACAGCAATTCGACAGAGCAGGTAATTCGCGCGGTTGTCCTGCTGAACTCC 195
 QY 525 TTAAGGCTTGTATTAAAGGACTATCAATTCGCCAATTGATGCTAATCAATCCGTC 584
 DB 196 ATGAGGCTTGTATATATGAGCTCCGCTCTACATGATGAAGCTGTGTAATCAATG 255
 QY 585 ACNAGAGAGGAGAGCAGAGAGAGTGTCTTAAGGCTCTTGTCTTACGTTTACGCTTT 644
 DB 256 ACAAGTGAAGGAAACAGAGAGATTTGCTGAGAGGTCACCTATATATGCTGTGCTG 315
 QY 645 CTTTCTCTGCGGTTTCTTCTGAGAGAGTCTCTATCGGTATGATCTGTTAGCAATG 704
 DB 316 CTGTTACAGCGTTTGTGTCTTGTGCGGTGATCCCATCGGAGTCGTCCTTGTGATG 375
 QY 705 ATGTGTGAGGAGATGATGATGATATATATGAGAGGTAAGTGGGTCACATAGATA 764
 DB 376 ATGAGCGGTGAGAGGTTTCCGTACATTTGTTGAGAGGATATGCTCAGCGAAGCTG 435
 QY 765 CTTTACAAACCAAGAAAGAGTTGGCAGAGACATCTCCATGTTATCTTCGCTTCTTC 824
 DB 436 CCATTCAATCGAAGAGAGGCTGGGCCGAGACATCTGATGTTTCTGTTTCTCG 495
 QY 825 ATCTTCATCGCATTAATTACTATTACTCAAGCTTTGGGTACCTTCAATGAACCTGGAA 884
 DB 496 CTGTCCGCGATGATGATGCTTACTTCTCAAGCTGGGTTACATGATGATTAATCGGAA 555
 QY 885 ACGACCTTGACAGAGTAGCAATGCTCAATGTCGCCAGGTAGTCGATCGCTACCC 944
 DB 556 GAGGACCTTGTAGAGCTGCGCTTGTGACATGACAGCAGATGATGAGTGCCTTCT 615
 QY 945 ATCACCAGATCAATTAAGACGACAAATTTTGGTTCTTGGCTACTAATTTAGCTGTTAT 1004
 DB 616 GTGACCGAAGTTGTATGATGACAAATATCTGTTCTTGGCCACCATGCTGTAGCTTTT 675
 QY 1005 TT 1006

DB 676 CT 677

Search completed: March 16, 2006, 16:17:26
 Job time : 1194 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: March 16, 2006, 18:50:48 ; Search time 90 Seconds
(without alignments)
1484.124 Million cell updates/sec

Title: US-10-634-548-2

Perfect score: 1536
Sequence: 1 MAATLPSPINHQCRCFGNN.....DDNISVPLATIAAYLSFGY 304

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378761 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 100%
Listing first 45 summaries

Database :

A_Geneseq_21:*

- 1: geneseqp1980s:*
- 2: geneseqp1990s:*
- 3: geneseqp2000s:*
- 4: geneseqp2001s:*
- 5: geneseqp2002s:*
- 6: geneseqp2003as:*
- 7: geneseqp2003bs:*
- 8: geneseqp2004s:*
- 9: geneseqp2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1536	100.0	304	8	ADJ98164
2	1536	100.0	315	8	ADJ98181
3	1523	99.2	301	8	ADJ98182
4	1523	99.2	304	3	AAG14596
5	1282	83.5	260	3	AAG14597
6	1266	82.4	304	8	ADJ98201
7	1266	82.4	319	8	ADJ98201
8	1140	74.2	226	8	ADJ98166
9	870	56.6	306	8	ADJ98171
10	864	56.2	298	8	ADJ98203
11	857.5	55.8	302	8	ADJ98207
12	791.5	51.5	314	8	ADJ98208
13	791.5	51.5	332	8	ADJ98121
14	791.5	51.5	657	8	ADJ98211
15	791.5	51.5	803	8	ADJ98210
16	786	51.2	309	8	ADJ98272
17	782.5	50.9	269	8	ADJ98209
18	777	50.6	303	8	ADJ98223
19	776	50.5	251	8	ADJ98203
20	773	50.3	300	8	ADJ98220
21	773	50.3	312	8	ADJ98197
22	752.5	49.0	267	8	ADJ98224
23	751	48.9	233	8	ADJ98167
24	751	48.9	236	8	ADJ98225

25	751	48.9	236	8	ADJ983050	Adx93050 Plant ful
26	718.5	46.8	292	8	ADJ98206	Adj98206 Soybean p
27	646	42.1	199	8	ADJ24930	Adj24930 Plant ful
28	643	41.9	129	3	AAG58667	Aag58667 Zea mays
29	615	40.0	191	8	ADJ98217	Adj98217 Sorghum p
30	607	39.5	259	8	ADJ58468	Adj58468 Plant pol
31	591	38.5	188	8	ADJ98222	Adj98222 Corn phyt
32	572.5	37.3	312	8	ADJ98227	Adj98227 Corn phyt
33	572.5	37.3	338	8	ADJ983092	Adx93092 Plant ful
34	565.5	36.8	246	8	ADJ98212	Adj98212 Rice phyt
35	562.5	36.6	302	8	ADJ98226	Adj98226 Corn phyt
36	552.5	36.0	296	8	ADJ98218	Adj98218 Bread whe
37	552.5	36.0	296	8	ADJ91976	Adx91976 Plant ful
38	544	35.4	346	8	ADT60333	Adt60333 Plant pol
39	542.5	35.3	319	8	ADJ98205	Adj98205 Soybean p
40	540.5	35.2	226	8	ADJ25323	Adj25323 Plant ful
41	539	35.1	288	8	ADJ98199	Adj98199 Leek phyt
42	539	35.1	289	8	ADJ94149	Adx94149 Plant ful
43	535.5	34.9	273	8	ADJ24874	Adj24874 Plant ful
44	530.5	34.5	307	3	AAG37869	Aag37869 Arabidops
45	530.5	34.5	307	8	ADJ98168	Adj98168 Thale cre

ALIGNMENTS

RESULT 1
ADJ98164
ID ADJ98164 standard; protein; 304 AA.
XX
AC ADJ98164;
XX
ADT 06-MAY-2004 (first entry)
XX
DE Thale cress LTT1 phytoL kinase wild-type protein.
XX
KM phytoL kinase; tocopherol biosynthesis; plant; drought resistance; LTT1;
KW thale cress; enzyme; wild-type.
XX
OS Arabidopsis thaliana.
XX
PN W02004013312-A2.
XX
PD 12-FEB-2004.
XX
PF 05-AUG-2003; 2003WO-US025276.
XX
PR 05-AUG-2002; 2002US-0400689P.
XX
PR 05-AUG-2003; 2003US-00634548.
XX
PA (MONS) MONSANTO TECHNOLOGY LLC.
XX
PI Norris SR, Lincoln K, Abad MS, Eilers R, Hartsuyker KK;
PI Hirschberg J, Karunananda B, Moshiri F, Stein JC, Valentin HE;
PI Venkatesh TV,
XX
XX WPI; 2004-157125/15.
XX
DR N-PSDB; ADJ98163.
XX
PT New phytoL kinase polynucleotides, useful in mediating tocopherol
PT biosynthesis and in producing plants with increased drought resistance.
XX
PS Claim 1; SEQ ID NO 2; 18pp; English.
XX
XX The invention relates to a novel substantially purified nucleic acid
XX molecule encoding a phytoL kinase. The nucleic acid molecules and
XX polypeptides of the invention may be useful in mediating tocopherol
XX biosynthesis and in producing plants with increased drought resistance.
XX The current sequence is that of the thale cress LTT1 phytoL kinase wild-
XX type protein of the invention.
XX
XX Sequence 304 AA;

Query Match 100.0%; Score 1536; DB 8; Length 304;
 Best Local Similarity 100.0%; Pred. No. 3,9e-162; Indels 0; Gaps 0;
 Matches 304; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAATPLSPINHQLCRFGNNSLTTHRFCSPEGLISSPCFGLTGMGSAATQLRARRSLISS 60
 DB 1 MAATPLSPINHQLCRFGNNSLTTHRFCSPEGLISSPCFGLTGMGSAATQLRARRSLISS 60
 QY 61 AATNSLHDVGAATVAAGVAVLSPESLTKNVIOQSRLVHLISGLFVLA MPI 120
 DB 61 AATNSLHDVGAATVAAGVAVLSPESLTKNVIOQSRLVHLISGLFVLA MPI 120
 QY 121 FSGSTEARYFAFVPLVNGRLVINGLSISPNMILISVTEGRAEELKGPLFVLAALL 180
 DB 121 FSGSTEARYFAFVPLVNGRLVINGLSISPNMILISVTEGRAEELKGPLFVLAALL 180
 QY 181 FSAVFWRPESPIGMIISLAMCGGDIADIMGRKFGSTKIPYNPKRSWAGSISMFTFGFFI 240
 DB 181 FSAVFWRPESPIGMIISLAMCGGDIADIMGRKFGSTKIPYNPKRSWAGSISMFTFGFFI 240
 QY 241 SIALLYSSSLGYLHMNETTLQRVAVSWAVTVESLPTDQDDNISVPLATILAAYL 300
 DB 241 SIALLYSSSLGYLHMNETTLQRVAVSWAVTVESLPTDQDDNISVPLATILAAYL 300
 QY 301 SFGY 304
 DB 301 SFGY 304

RESULT 2
 ADX91881
 ID ADX91881 standard; protein; 315 AA.

XX AC ADX91881;
 XX DT 21-APR-2005 (first entry)
 XX DE Plant full length insert polypeptide seqid 54545.
 XX DB plant protectant; plant growth regulant; gene therapy; plant;
 KW recombinant DNA construct; physical array; plant breeding marker;
 KW cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;
 KW extreme osmotic condition; pathogen tolerance; pest tolerance;
 KW growth rate; cell cycle pathway; disease resistance;
 KW galactomannan production; lignin production; plant growth regulator;
 KW yield; plant growth; plant development; seed oil; protein yield;
 KW protein content.
 XX OS Unidentified.
 XX PN US2004034888-A1.
 XX PD 19-FEB-2004.
 XX PF 28-APR-2003; 2003US-00425114.
 XX PR 06-MAY-1999; 99US-00304517.
 XX RR 05-NOV-2001; 2001US-00985678.
 XX PA (LIU/J.) LIU J.
 XX PA (ZHOU/) ZHOU Y.
 XX PA (KOVA/) KOVALLIC D K.
 XX PA (SCRE/) SCREEN S E.
 XX PA (TABAK/) TABASKA J E.
 XX PA (CAOY/) CAO Y.
 XX PI Liu J, Zhou Y, Kovallik DK, Screen SE, Tabaska JE, Cao Y;
 XX DE MPI; 2004-180133/17.
 XX PT New recombinant DNA construct, useful for improving plant tolerance to
 PT cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or
 PT pests, for conferring increased resistance to plant disease, or for

PT improving yield.

XX Claim 1; SEQ ID NO 54545; 15pp; English.

XX The invention describes a recombinant DNA construct comprising a
 CC polynucleotide consisting of a sequence encoding an amino acid sequence
 CC available in electronic form from the US patent office at
 CC ftp.segdata.uspto.gov/sequence.html; DocId:2004034888. The polynucleotide
 CC of the invention are also useful in physical arrays of molecules and as
 CC plant breeding markers. The recombinant DNA construct is useful for
 CC improving plant tolerance to cold, heat, drought, herbicides, extreme
 CC osmotic conditions, pathogens or pests, for manipulating growth rate in
 CC plant cells by modification of the cell cycle pathway, for conferring
 CC increased resistance to plant disease, for producing galactomannan,
 CC lignin or plant growth regulator, for increasing the rate of homologous
 CC recombination in plants, for improving yield by modification of
 CC photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake
 CC or by providing improved plant growth and development under at least one
 CC stress condition or for modifying seed oil or protein yield and/or
 CC content. This is the amino acid sequence of a plant full length insert
 CC polypeptide that can be used in the recombinant DNA construct of the
 CC invention.

XX Sequence 315 AA;

Query Match 100.0%; Score 1536; DB 8; Length 315;
 Best Local Similarity 100.0%; Pred. No. 4.1e-162;
 Matches 304; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAATPLSPINHQLCRFGNNSLTTHRFCSPEGLISSPCFGLTGMGSAATQLRARRSLISS 60
 DB 12 MAATPLSPINHQLCRFGNNSLTTHRFCSPEGLISSPCFGLTGMGSAATQLRARRSLISS 71
 QY 61 AATNSLHDVGAATVAAGVAVLSPESLTKNVIOQSRLVHLISGLFVLA MPI 120
 DB 72 AATNSLHDVGAATVAAGVAVLSPESLTKNVIOQSRLVHLISGLFVLA MPI 131
 QY 121 FSGSTEARYFAFVPLVNGRLVINGLSISPNMILISVTEGRAEELKGPLFVLAALL 180
 DB 132 FSGSTEARYFAFVPLVNGRLVINGLSISPNMILISVTEGRAEELKGPLFVLAALL 191
 QY 181 FSAVFWRPESPIGMIISLAMCGGDIADIMGRKFGSTKIPYNPKRSWAGSISMFTFGFFI 240
 DB 192 FSAVFWRPESPIGMIISLAMCGGDIADIMGRKFGSTKIPYNPKRSWAGSISMFTFGFFI 251
 QY 241 SIALLYSSSLGYLHMNETTLQRVAVSWAVTVESLPTDQDDNISVPLATILAAYL 300
 DB 252 SIALLYSSSLGYLHMNETTLQRVAVSWAVTVESLPTDQDDNISVPLATILAAYL 311
 QY 301 SFGY 304
 DB 312 SFGY 315

RESULT 3
 ADX91882

ID ADX91882 standard; protein; 301 AA.

XX AC ADX91882;
 XX DT 21-APR-2005 (first entry)
 XX DE Plant full length insert polypeptide seqid 54546.
 XX DB plant protectant; plant growth regulant; gene therapy; plant;
 KW recombinant DNA construct; physical array; plant breeding marker;
 KW cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;
 KW extreme osmotic condition; pathogen tolerance; pest tolerance;
 KW growth rate; cell cycle pathway; disease resistance;
 KW galactomannan production; lignin production; plant growth regulator;
 KW yield; plant growth; plant development; seed oil; protein yield;
 KW protein content.

OS Unidentified.
XX US200403488-A1.
XX 19-FEB-2004.
XX 28-APR-2003; 2003US-00425114.
XX 06-MAY-1999; 99US-00304517.
XX 05-NOV-2001; 2001US-00985678.
XX (LIU/J) LIU J.
XX (ZHOU/Y) ZHOU Y.
XX (KOVAN/) KOVALIC D K.
XX (SCRE/) SCREEN S E.
XX (TABAKA/) TABASKA J E.
XX (CAO/Y) CAO Y.
XX Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JE, Cao Y;
PI WPI; 2004-180133/17.
XX New recombinant DNA construct, useful for improving plant tolerance to
PT cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or
PT pests, for conferring increased resistance to plant disease, or for
PT improving yield.
XX Claim 1; SEQ ID NO 54546; 15bp; English.
XX The invention describes a recombinant DNA construct comprising a
CC polynucleotide consisting of a sequence encoding an amino acid sequence
CC available in electronic form from the US patent office at
CC ftp.seqdata.uspto.gov/sequence.html?docID:200403488. The polynucleotide
CC of the invention are also useful in physical arrays of molecules and as
CC plant breeding markers. The recombinant DNA construct is useful for
CC improving plant tolerance to cold, heat, drought, herbicides, extreme
CC osmotic conditions, pathogens or pests, for manipulating growth rate in
CC plant cells by modification of the cell cycle pathway, for conferring
CC increased resistance to plant disease, for producing galactomannan,
CC lignin or plant growth regulators, for increasing the rate of homologous
CC recombination in plants, for improving yield by modification of
CC photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake
CC or by providing improved plant growth and development under at least one
CC stress condition or for modifying seed oil or protein yield and/or
CC content. This is the amino acid sequence of a plant full length insert
CC polypeptide that can be used in the recombinant DNA construct of the
CC invention.
XX Sequence 301 AA;
SQ
Query Match 99.2%; Score 1523; DB 8; Length 301;
Best Local Similarity 100.0%; Pred. No. 1.1e-160;
Matches 301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 4 TLPLSPINHOICRFENNSLTTHRFCSPPFLISPPFGLTGMGSAFOLARRRSIISSAVA 63
DB 1 TLPLSPINHOICRFENNSLTTHRFCSPPFLISPPFGLTGMGSAFOLARRRSIISSAVA 60
QY 64 TNSLHDVGATVAVALGAYVALVSFESLTGRNVIOQSLSRKLVHLSGLFVLAAMPISFG 123
DB 61 TNSLHDVGATVAVALGAYVALVSFESLTGRNVIOQSLSRKLVHLSGLFVLAAMPISFG 120
QY 124 STEARYFAFVPLVNGRLVINGLSISPNMSLKSVTREGRAEELKGPFLFYVALALFSA 183
DB 121 STEARYFAFVPLVNGRLVINGLSISPNMSLKSVTREGRAEELKGPFLFYVALALFSA 180
QY 184 VFPPRESPIGMIISLMMCGGDIIMGRKFGSTIPIVNPBKSNAGSISMPFSGFISIA 243
DB 181 VFPPRESPIGMIISLMMCGGDIIMGRKFGSTIPIVNPBKSNAGSISMPFSGFISIA 240
QY 244 LLYYSSIGYIHMNMETTLQRYAVMSVATVSESIPITDQDDNISVPLATILAAVLSFG 303
DB 241 LLYYSSIGYIHMNMETTLQRYAVMSVATVSESIPITDQDDNISVPLATILAAVLSFG 300

QY 304 Y 304
DB 301 Y 301
RESULT 4
ID AAG14596 standard; protein; 304 AA.
XX AAG14596;
XX 17-OCT-2000 (first entry)
XX Arabidopsis thaliana protein fragment SEQ ID NO: 14520.
XX Protein identification; signal transduction pathway; metabolic pathway;
XX hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
XX Arabidopsis thaliana.
OS EP1033405-A2.
XX EP1033405-A2.
XX 06-SEP-2000.
XX 25-FEB-2000; 2000EP-00301439.
XX 25-FEB-1999; 99US-0121825P.
XX 05-MAR-1999; 99US-0123180P.
XX 09-MAR-1999; 99US-0123548P.
XX 23-MAR-1999; 99US-0125788P.
XX 25-MAR-1999; 99US-0126264P.
XX 29-MAR-1999; 99US-0126785P.
XX 01-APR-1999; 99US-0127462P.
XX 06-APR-1999; 99US-0128234P.
XX 08-APR-1999; 99US-0128714P.
XX 16-APR-1999; 99US-0129845P.
XX 19-APR-1999; 99US-0130077P.
XX 21-APR-1999; 99US-0130449P.
XX 23-APR-1999; 99US-0130510P.
XX 28-APR-1999; 99US-0130891P.
XX 30-APR-1999; 99US-0131449P.
XX 30-APR-1999; 99US-0132048P.
XX 30-APR-1999; 99US-0132407P.
XX 04-MAY-1999; 99US-0132484P.
XX 05-MAY-1999; 99US-0132485P.
XX 06-MAY-1999; 99US-0132486P.
XX 07-MAY-1999; 99US-0132863P.
XX 11-MAY-1999; 99US-0134256P.
XX 14-MAY-1999; 99US-0134218P.
XX 14-MAY-1999; 99US-0134219P.
XX 14-MAY-1999; 99US-0134221P.
XX 14-MAY-1999; 99US-0134370P.
XX 18-MAY-1999; 99US-0134768P.
XX 19-MAY-1999; 99US-0134941P.
XX 20-MAY-1999; 99US-0135124P.
XX 21-MAY-1999; 99US-0135333P.
XX 24-MAY-1999; 99US-0135629P.
XX 25-MAY-1999; 99US-0136021P.
XX 27-MAY-1999; 99US-0136392P.
XX 28-MAY-1999; 99US-0136782P.
XX 01-JUN-1999; 99US-0137222P.
XX 03-JUN-1999; 99US-0137528P.
XX 04-JUN-1999; 99US-0137502P.
XX 07-JUN-1999; 99US-0137724P.
XX 08-JUN-1999; 99US-0138094P.
XX 10-JUN-1999; 99US-0138540P.
XX 10-JUN-1999; 99US-0138847P.
XX 14-JUN-1999; 99US-0139119P.
XX 16-JUN-1999; 99US-0139452P.
XX 16-JUN-1999; 99US-0139453P.

PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139751P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144336P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144633P.
PR 21-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145086P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 06-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.

PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 99.7%; Score 1523; DB 3; Length 304;
Best Local Similarity 99.7%; Pred. No. 1,1e-160;
Matches 300; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAATPLSPINHQCRCFNNSLTTHRFCSPGFLISSPCFGLTGMGSAIOLRARRSLISS 60
DB 1 MDATPLSPINHQCRCFNNSLTTHRFCSPGFLISSPCFGLTGMGSAIOLRARRSLISS 60
QY 61 AVATNSLIADVGATVAVIGGAYALVLSFESLTKRNVIOOSLSRKLWHLISGLPLFLAPI 120

Db 61 AVATNSLHDVGTAVVAGVAYLVLSPESTLKRNVQGSRLKVLHLSGLFVLA MPI 120
Qy 121 FSGSTEARYPAFVPLVNGLRVINGLSISPSMTLKSYTRGRARELLKGFPLVYAL 180
Db 121 FSGSTEARYPAFVPLVNGLRVINGLSISPSMTLKSYTRGRARELLKGFPLVYAL 180
Qy 181 FSAVFMEESPFGMTSLAMMCGGDIADIMGRKFGSTKIPYAPRKSWAGSISMPFEGPFI 240
Db 181 FSAVFMEESPFGMTSLAMMCGGDIADIMGRKFGSTKIPYAPRKSWAGSISMPFEGPFI 240
Qy 241 SIALLYYSSLSGLYHNMWETLQRVAVMSVATVVESSLPTDQLDNISVPLATIIAAYL 300
Db 241 SIALLYYSSLSGLYHNMWETLQRVAVMSVATVVESSLPTDQLDNISVPLATIIAAYL 300
Qy 301 SFGY 304
Db 301 SFGY 304

RESULT 5
AAG14597
ID AAG14597 standard; protein; 260 AA.
XX AAG14597;
AC
XX 17-OCT-2000 (first entry)
DT
XX Arabidopsis thaliana protein fragment SEQ ID NO: 14521.
DE
XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX Arabidopsis thaliana.
OS
XX EP1033405-A2.
FN
XX 06-SEP-2000.
PD
XX 25-FEB-2000; 2000EP-00301439.
PF
XX
XX 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0128845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130549P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132485P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.

PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 28-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.

PR 03-AUG-1999; 99US-0147038P.
 PR 04-AUG-1999; 99US-0147204P.
 PR 05-AUG-1999; 99US-0147302P.
 PR 05-AUG-1999; 99US-0147192P.
 PR 05-AUG-1999; 99US-0147260P.
 PR 06-AUG-1999; 99US-0147303P.
 PR 06-AUG-1999; 99US-0147416P.
 PR 09-AUG-1999; 99US-0147933P.
 PR 10-AUG-1999; 99US-0148171P.
 PR 11-AUG-1999; 99US-0148312P.
 PR 12-AUG-1999; 99US-0148341P.
 PR 13-AUG-1999; 99US-0148565P.
 PR 16-AUG-1999; 99US-0149368P.
 PR 17-AUG-1999; 99US-0149175P.
 PR 18-AUG-1999; 99US-0149426P.
 PR 20-AUG-1999; 99US-0149722P.
 PR 20-AUG-1999; 99US-0149723P.
 PR 20-AUG-1999; 99US-0149923P.
 PR 23-AUG-1999; 99US-0149902P.
 PR 23-AUG-1999; 99US-0149930P.
 PR 25-AUG-1999; 99US-0150566P.
 PR 26-AUG-1999; 99US-0150884P.
 PR 27-AUG-1999; 99US-0151065P.
 PR 27-AUG-1999; 99US-0151066P.
 PR 27-AUG-1999; 99US-0151080P.
 PR 30-AUG-1999; 99US-0151303P.
 PR 31-AUG-1999; 99US-0151438P.
 PR 01-SEP-1999; 99US-0151930P.
 PR 10-SEP-1999; 99US-0152363P.
 PR 13-SEP-1999; 99US-0153070P.
 PR 15-SEP-1999; 99US-0154018P.
 PR 16-SEP-1999; 99US-0154032P.
 PR 20-SEP-1999; 99US-0154779P.
 PR 22-SEP-1999; 99US-0155139P.
 PR 23-SEP-1999; 99US-0155486P.
 PR 24-SEP-1999; 99US-0155659P.
 PR 28-SEP-1999; 99US-0156458P.
 PR 29-SEP-1999; 99US-0156596P.
 PR 04-OCT-1999; 99US-0157117P.
 PR 05-OCT-1999; 99US-0157753P.
 PR 06-OCT-1999; 99US-0157865P.
 PR 07-OCT-1999; 99US-0158023P.
 PR 08-OCT-1999; 99US-0158232P.
 PR 12-OCT-1999; 99US-0158363P.
 PR 13-OCT-1999; 99US-0159293P.
 PR 13-OCT-1999; 99US-0159294P.
 PR 13-OCT-1999; 99US-0159295P.
 PR 14-OCT-1999; 99US-0159329P.
 PR 14-OCT-1999; 99US-0159330P.
 PR 14-OCT-1999; 99US-0159331P.
 PR 14-OCT-1999; 99US-0159637P.
 PR 14-OCT-1999; 99US-0159638P.
 PR 18-OCT-1999; 99US-0159584P.
 PR 21-OCT-1999; 99US-0160741P.
 PR 21-OCT-1999; 99US-0160767P.
 PR 21-OCT-1999; 99US-0160768P.
 PR 21-OCT-1999; 99US-0160770P.
 PR 21-OCT-1999; 99US-0160814P.
 PR 21-OCT-1999; 99US-0160815P.
 PR 22-OCT-1999; 99US-0160980P.
 PR 22-OCT-1999; 99US-0160981P.
 PR 22-OCT-1999; 99US-0160989P.
 PR 25-OCT-1999; 99US-0161404P.
 PR 25-OCT-1999; 99US-0161405P.
 PR 25-OCT-1999; 99US-0161406P.
 PR 26-OCT-1999; 99US-0161359P.
 PR 26-OCT-1999; 99US-0161360P.
 PR 26-OCT-1999; 99US-0161361P.
 PR 28-OCT-1999; 99US-0161920P.
 PR 28-OCT-1999; 99US-0161992P.

PR 28-OCT-1999; 99US-0161993P.
 PR 29-OCT-1999; 99US-0162142P.

Query Match 83.5%; Score 1282; DB 3; Length 260;
 Best Local Similarity 98.5%; Pred. No. 6.8e-134;
 Matches 256; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 45 MGSATQRLARRSLISSAVATNTSLHDVGATVAVLGAVALVLSFESLTQRNVIOQSLSRK 104
 DB 1 MGSATQRLARRSLISSAVATNTSLHDVGATVAVLGAVALVLSFESLTQRNVIOQSLSRK 60
 QY 105 LVHTLSGLFLVLAAMPFGSGTEARYFAFPVPLVNGGLVINGLSISPSNMLIKSVTEGR 164
 DB 61 LVHTLSGLFLVLAAMPFGSGTEARYFAFPVPLVNGGLVINGLSISPSNMLIKSVTEGR 120
 QY 165 ABEILKGPLFYVLAALSAVFPWRSPFGIMISLMMCGDGIADIMGRKFGSTKIPYPR 224
 DB 121 ABEILKGPLFYVLAALSAVFPWRSPFGIMISLMMCGDGIADIMGRKFGSTKIPYPR 180
 QY 225 KSMAGSISMFIKGFIFISIALLYYSSLGYLHMMNETTLQRYAVMSVATVVESSLPITDQL 284
 DB 181 KSMAGSISMFIKGFIFISIALLYYSSLGYLHMMNETTLQRYAVMSVATVVESSLPITDQL 240
 QY 285 DDNISVPLATTLAAYLSFGY 304
 DB 241 DDNISVPLATTLAAYLSFGY 260

RESULT 6
 ADJ98201
 ID ADJ98201 standard; protein; 304 AA.

AC ADJ98201;

DT 06-MAY-2004 (first entry)

XX Rape phytol kinase protein.

XX phytol kinase; tocopherol biosynthesis; plant; drought resistance;

KW enzyme; rape.

OS Brassica napus.

PN WO2004013312-A2.

PD 12-FEB-2004.

PF 05-AUG-2003; 2003WO-US025276.

PR 05-AUG-2002; 2002US-0400689P.

PR 05-AUG-2003; 2003US-00634548.

XX (MONS) MONSANTO TECHNOLOGY LLC.

XX Norris SR, Lincoln K, Abad MS, Eilers R, Hartuysen KK;

PI Hirschberg J, Karunanandaa B, Moshiri F, Stein JC, Valentin HE;

PI Venkatesh TV;

XX WPI; 2004-157125/15.

PT New phytol kinase polynucleotides, useful in mediating tocopherol
 biosynthesis and in producing plants with increased drought resistance.

PS Claim 1; SEQ ID NO 39; 189pp; English.

CC The invention relates to a novel substantially purified nucleic acid
 CC molecule encoding a phytol kinase. The nucleic acid molecules and
 CC polypeptides of the invention may be useful in mediating tocopherol
 CC biosynthesis and in producing plants with increased drought resistance.
 CC The current sequence is that of a phytol kinase protein of the invention.

XX Sequence 304 AA;

SQ

Query Match 82.4%; Score 1266; DB 8; Length 304;
 Best Local Similarity 81.4%; Pred. No. 5.2e-133;
 Matches 250; Conservative 24; Mismatches 27; Indels 6; Gaps 2;

QY 1 MAATLPSPINHQCRRFGNN---SLTTRPFCSPGFLISSPCFGLTGMGSAATQLRARSL 57
 DB 1 MAALPLSPVSHQCRRISNRFWYAMTPRCPSP---VSPCYIGVKIGSSQLRARHPL 57
 QY 58 ISSAVATNSLLHDVGATVAVALGAYALVLSFESLTGRNVIOQSLSRKLVHLSGLLFYLA 117
 DB 58 ISSASTDYLLHDVGATVAVALGAYALVLSFESLTGRNVIOQSLSRKLVHLSGLLFALS 117
 QY 118 WPIRSGSTEARYPFAFVPLVNGRLVINGLSISNSMLIKSTREGRAEELKGPLFYVL 177
 DB 118 WPIRSGSTEARYPFAFVPLVNGRLVINGLSISNSMLIKSTREGRAEELKGPLFYVL 177
 QY 178 ALPSAVFPMRSPGIMISLAMCGGDIADIMGRKFGSTKIPYNPGRKSWAGSISMFIYG 237
 DB 178 ALVAAVFPWRDSTGIMISLAMCGGDIADIMGRKFGSTKIPYNPGRKSLAGSISMFIYG 237
 QY 238 FFISIALLYYSSLSGYLHMNMETTLQRYAMVSMVATVYESLPITDQDDNISVPLATTLA 297
 DB 238 FFISIGLLYYSSLSGYLHMNMETTFTRVAIVSLVATIVESLPITDQDDNISVPLATTLA 297
 QY 298 AYLSFGY 304
 DB 298 AYLSFGY 304

RESULT 7
 ADX91966
 ID ADX91966 standard; protein, 319 AA.

XX AC ADX91966;
 XX DT 21-Apr-2005 (first entry)
 XX DE Plant full length insert polypeptide seqid 54630.
 XX KW plant protectant; plant growth regulant; gene therapy; plant;
 KW recombinant DNA construct; physical array; plant breeding marker;
 KW cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;
 KW extreme osmotic condition; pathogen tolerance; pest tolerance;
 KW growth rate; cell cycle pathway; disease resistance;
 KW galactomannan production; lignin production; plant growth regulator;
 KW yield; plant growth; plant development; seed oil; protein yield;
 KW protein content.
 XX OS unidentified.
 XX PA US2004034888-A1.
 XX PD 19-FEB-2004.
 XX PF 28-APR-2003; 2003US-00425114.
 XX PR 06-MAY-1999; 99US-00304517.
 XX PR 05-NOV-2001; 2001US-00985678.
 XX PA (LITUU/) LIU J.
 XX PA (ZHOU/) ZHOU Y.
 XX PA (KOVA/) KOVALIC D K.
 XX PA (SCRE/) SCREEN S E.
 XX PA (TABAKA/) TABAKA J E.
 XX PA (CAOY/) CAO Y.
 XX PI Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaka JB, Cao Y;
 XX DR WPI; 2004-180133/17.
 XX PT New recombinant DNA construct, useful for improving plant tolerance to
 PT cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or
 PT pests, for conferring increased resistance to plant disease, or for

PT improving yield.

XX Claim 1; SEQ ID NO 54630; 15bp; English.

XX The invention describes a recombinant DNA construct comprising a
 CC polynucleotide consisting of a sequence encoding an amino acid sequence
 CC available in electronic form from the US patent office at
 CC ftp.segdata.uspco.gov/sequence.html?docid:2004034888. The polynucleotide
 CC of the invention are also useful in physical arrays of molecules and as
 CC plant breeding markers. The recombinant DNA construct is useful for
 CC improving plant tolerance to cold, heat, drought, herbicides, extreme
 CC osmotic conditions, pathogens or pests, for manipulating growth rate in
 CC plant cells by modification of the cell cycle pathway, for conferring
 CC increased resistance to plant disease, for producing galactomannan,
 CC lignin or plant growth regulators, for increasing the rate of homologous
 CC recombination in plants, for improving yield by modification of
 CC photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake
 CC or by providing improved plant growth and development under at least one
 CC stress condition or for modifying seed oil or protein yield and/or
 CC content. This is the amino acid sequence of a plant full length insert
 CC polypeptide that can be used in the recombinant DNA construct of the
 CC invention.

XX Sequence 319 AA;

Query Match 82.4%; Score 1266; DB 8; Length 319;
 Best Local Similarity 81.4%; Pred. No. 5.5e-132;
 Matches 250; Conservative 24; Mismatches 27; Indels 6; Gaps 2;

QY 1 MAATLPSPINHQCRRFGNN---SLTTRPFCSPGFLISSPCFGLTGMGSAATQLRARSL 57
 DB 16 MAALPLSPVSHQCRRISNRFWYAMTPRCPSP---VSPCYIGVKIGSSQLRARHPL 72
 QY 58 ISSAVATNSLLHDVGATVAVALGAYALVLSFESLTGRNVIOQSLSRKLVHLSGLLFYLA 117
 DB 73 ISSASTDYLLHDVGATVAVALGAYALVLSFESLTGRNVIOQSLSRKLVHLSGLLFALS 132
 QY 118 WPIRSGSTEARYPFAFVPLVNGRLVINGLSISNSMLIKSTREGRAEELKGPLFYVL 177
 DB 133 WPIRSGSTEARYPFAFVPLVNGRLVINGLSISNSMLIKSTREGRAEELKGPLFYVL 192
 QY 178 ALPSAVFPMRSPGIMISLAMCGGDIADIMGRKFGSTKIPYNPGRKSWAGSISMFIYG 237
 DB 193 ALVAAVFPWRDSTGIMISLAMCGGDIADIMGRKFGSTKIPYNPGRKSLAGSISMFIYG 252
 QY 238 FFISIALLYYSSLSGYLHMNMETTLQRYAMVSMVATVYESLPITDQDDNISVPLATTLA 297
 DB 253 FFISIGLLYYSSLSGYLHMNMETTFTRVAIVSLVATIVESLPITDQDDNISVPLATTLA 312
 QY 298 AYLSFGY 304
 DB 313 AYLSFGY 319

RESULT 8
 ADX98166
 ID ADX98166 standard; protein, 226 AA.

XX AC ADX98166;
 XX DT 06-MAY-2004 (first entry)
 XX DE Thale cress LTT1 phyto kinase mutant protein.
 XX KW phyto kinase; tocopherol biosynthesis; plant; drought resistance; LTT1;
 KW thale cress; enzyme; mutant; mutein.
 XX OS Arabidopsis thaliana.
 XX PA Key Location/Qualifiers
 XX FT Misc-difference 226
 XX FT /note="Wild-type Ser, Trp is substituted for Ser, STOP"

```

PN WO2004013312-A2.
XX
PD 12-FEB-2004.
XX
XX 05-AUG-2003; 2003WO-US025276.
XX
XX 05-AUG-2002; 2002US-0400689P.
XX
XX 05-AUG-2003; 2003US-00634548.
XX
XX (MONS ) MONSANTO TECHNOLOGY LLC.
XX
XX Norris SR, Lincoln K, Abad MS, Eljers R, Harteuwerker KK,
XX Hirschberg J, Karunanandaa B, Moshiri F, Stein JC, Valentin HR,
XX Venkatesh TV;
XX
XX WPI; 2004-157125/15.
XX
XX N-PSDB; ADJ98165.
XX
XX New phyto kinase polynucleotides, useful in mediating tocopherol
XX biosynthesis and in producing plants with increased drought resistance.
XX
XX Example 2; SEQ ID NO 4; 189pp; English.
XX
XX The invention relates to a novel substantially purified nucleic acid
XX molecule encoding a phyto kinase. The nucleic acid molecules and
XX polypeptides of the invention may be useful in mediating tocopherol
XX biosynthesis and in producing plants with increased drought resistance.
XX The current sequence is that of the thale cress LTT1 phyto kinase mutant
XX protein of the invention.
XX
XX Sequence 226 AA;
XX
XX Query Match 74.2%; Score 1140; DB 8; Length 226;
XX Best Local Similarity 100.0%; Pred. No. 3.8e-118;
XX Matches 226; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 MAATPLSPINHQLCRFGNNLSITTRFCSPGFLISSPCFGLTGWGSAATOLRARRSLISS 60
XX 1 MAATPLSPINHQLCRFGNNLSITTRFCSPGFLISSPCFGLTGWGSAATOLRARRSLISS 60
XX
XX 61 AVATNSLHDVAGATVAIVGAYALVLSFESLTKRNVIOQSLSRKVHLISGLLPYLA 120
XX 61 AVATNSLHDVAGATVAIVGAYALVLSFESLTKRNVIOQSLSRKVHLISGLLPYLA 120
XX
XX 121 FSGSTEARYPFAFVPLVNGRLVINGLSISPSNMLIKSVTRGRABEILKGPLFYVAL 180
XX 121 FSGSTEARYPFAFVPLVNGRLVINGLSISPSNMLIKSVTRGRABEILKGPLFYVAL 180
XX
XX 181 FSAVFPWRESPIGMSLMMCGGDIADIMGRKFGSTKIPYVPRKS 226
XX 181 FSAVFPWRESPIGMSLMMCGGDIADIMGRKFGSTKIPYVPRKS 226
XX
XX RESULT 9
XX ADX91971
XX ID ADX91971 standard; protein; 306 AA.
XX
XX AC ADX91971;
XX
XX DT 21-APR-2005 (first entry)
XX
XX DE Plant full length insert polypeptide seqid 54635.
XX
XX KW plant protectant; plant growth regulant; gene therapy; plant;
XX recombinant DNA construct; physical array; plant breeding marker;
XX cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;
XX extreme osmotic condition; pathogen tolerance; pest tolerance;
XX growth rate; cell cycle pathway; disease resistance;
XX galactomanan production; lignin production; plant growth regulator;
XX yield; plant growth; plant development; seed oil; protein yield;
XX protein content.
XX
XX OS unidentified.

```

```

XX US2004034888-A1.
XX
XX 19-FEB-2004.
XX
XX 28-APR-2003; 2003US-00425114.
XX
XX 06-MAY-1999; 99US-00304517.
XX
XX 05-NOV-2001; 2001US-00985678.
XX
XX (LITUJ/) LIU J.
XX
XX (ZHOU/) ZHOU Y.
XX
XX (KOVA/) KOVALIC D K.
XX
XX (SCRE/) SCREEN S E.
XX
XX (TABAS/) TABASKA J E.
XX
XX (CAO/) CAO Y.
XX
XX Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JE, Cao Y;
XX
XX WPI; 2004-180133/17.
XX
XX New recombinant DNA construct, useful for improving plant tolerance to
XX cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or
XX pests, for conferring increased resistance to plant disease, or for
XX improving yield.
XX
XX Claim 1; SEQ ID NO 54635; 15pp; English.
XX
XX The invention describes a recombinant DNA construct comprising a
XX polynucleotide consisting of a sequence encoding an amino acid sequence
XX available in electronic form from the US patent office at
XX ftp.segdata.uspto.gov/sequence.html?docID:200403488. The polynucleotide
XX of the invention are also useful in physical arrays of molecules and as
XX plant breeding markers. The recombinant DNA construct is useful for
XX improving plant tolerance to cold, heat, drought, herbicides, extreme
XX osmotic conditions, pathogens or pests, for manipulating growth rate in
XX plant cells by modification of the cell cycle pathway, for conferring
XX increased resistance to plant disease, for producing galactomanan,
XX lignin or plant growth regulators, for increasing the rate of homologous
XX recombination in plants, for improving yield by modification of
XX photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake
XX or by providing improved plant growth and development under at least one
XX stress condition or for modifying seed oil or protein yield and/or
XX content. This is the amino acid sequence of a plant full length insert
XX polypeptide that can be used in the recombinant DNA construct of the
XX invention.
XX
XX Sequence 306 AA;
XX
XX Query Match 56.6%; Score 870; DB 8; Length 306;
XX Best Local Similarity 56.7%; Pred. No. 7.8e-88;
XX Matches 174; Conservative 51; Mismatches 72; Indels 10; Gaps 2;
XX
XX 2 AATPLSPINHQLCRFGNNLSITTRFCS---PGFLISSPCFGLTGWGSAATOLRARRSL 57
XX 6 AATMSLS-----LSFTPIILSRHYSAVFPFPPRLFLSPILPTTSRPPILYRADORATA 59
XX
XX 58 ISSAVATNSLHDVAGATVAIVGAYALVLSFESLTKRNVIOQSLSRKVHLISGLLPYLA 117
XX 60 LSATAVTASIFRDRNASASVAGAYALVFTFDITOKELIQNLSRKVHLISGLLPYLA 119
XX
XX 118 WPIFSGSTEARYPFAFVPLVNGRLVINGLSISPSNMLIKSVTRGRABEILKGPLFYVAL 177
XX 120 WPIFSGSTEARYPFAFVPLVNGRLVINGLSISPSNMLIKSVTRGRABEILKGPLFYVAL 179
XX
XX 178 ALBSAVFPWRESPIGMSLMMCGGDIADIMGRKFGSTKIPYVPRKSMSGSI 237
XX 180 MLCMCALVFPWRESPIGMSLMMCGGDIADIMGRKFGSTKIPYVPRKSMSGSI 239
XX
XX 238 FFISIALLYYSSLSGLYHMMWETTLQRYAMTSMVATVESPITDQLDNISVPLATIIA 297
XX 240 FTISIGMLYYTSALQYLDMDGYTLHRVAFISIVATVESPISMLIDNISVPLASMLA 299

```

Oy 298 AYISFGY 304
 |||:
 Db 300 AYLTFGH 306

RESULT 10

ID ADJ98203 standard; protein; 298 AA.

XX ADJ98203;

XX 06-MAY-2004 (first entry)

XX Upland cotton phytochrome kinase protein.

XX phytochrome kinase; tocopherol biosynthesis; plant; drought resistance;

XX enzyme; upland cotton.

XX Gossypium hirsutum.

XX MO2004013312-A2.

XX 12-FEB-2004.

XX 05-AUG-2003; 2003WO-US025276.

XX 05-AUG-2002; 2002US-0400689P.

XX 05-AUG-2003; 2003US-00634548.

XX (MONS) MONSANTO TECHNOLOGY LLC.

XX Norris SR, Lincoln K, Abad MS, Eilers R, Hartuysen KK;

XX Hirschberg J, Karunanandaa B, Moshiri F, Stein JC, Valentin HE;

XX Venkatesh TV;

XX WPI; 2004-157125/15.

XX New phytochrome kinase polynucleotides, useful in mediating tocopherol

XX biosynthesis and in producing plants with increased drought resistance.

XX Claim 1; SEQ ID NO 41; 189pp; English.

XX The invention relates to a novel substantially purified nucleic acid

XX molecule encoding a phytochrome kinase. The nucleic acid molecules and

XX polypeptides of the invention may be useful in mediating tocopherol

XX biosynthesis and in producing plants with increased drought resistance.

XX The current sequence is that of a phytochrome kinase protein of the invention.

XX Sequence 298 AA;

XX Query Match 56.2%; Score 864; DB 8; Length 298;

XX Best Local Similarity 60.1%; Pred. No. 3.5e-87;

XX Matches 167; Conservative 46; Mismatches 65; Indels 0; Gaps 0;

Oy 27 FCSFGFLSSPCFGLTGMGATOLRARRSLISSAVATNSLHDVGAATVAVGAYALV 86

Db 21 PPPPFLSLSPILPTTSRFPILYAPQATASATVATASITRDPASASVPAAGAYALV 80

Oy 87 SFESLTKRNVIOQSLRSKLVHLSGLFLVAMPFSGSTEARYPFAVPLVNGSLTVNG 146

Db 81 TFDILTOGRELQONSRKLVHLSGLFLVAMPFSGSTEARYPFAVPLVNGSLTVNG 140

Oy 147 LSISSNMLIKSVTRREGABELKGPLFYVALALSAVFPFRESITGMSLMMCGGDCI 206

Db 141 LSLTDQSLIKSVTRREGABELKGPLFYVALALSAVFPFRESITGMSLMMCGGDCI 200

Oy 207 ADIMRKGSTKIPYNPKNAGSISMPFISFISALTYSSLSGYLHMWETTLQRYA 266

Db 201 ADIIRKKGSSKIPYNPKNAGSISMPFISFISALTYSSLSGYLHMWETTLQRYA 260

Oy 267 MVSNAVATVESLPTDQDDNISVPLATITLAAYISFGY 304

Db 261 FISTVATVESLPTDQDDNISVPLATITLAAYISFGY 298

RESULT 11

ID ADJ98207 standard; protein; 302 AA.

XX ADJ98207;

XX 06-MAY-2004 (first entry)

XX Soybean phytochrome kinase protein.

XX phytochrome kinase; tocopherol biosynthesis; plant; drought resistance;

XX enzyme; soybean.

XX Glycine max.

XX MO2004013312-A2.

XX 12-FEB-2004.

XX 05-AUG-2003; 2003WO-US025276.

XX 05-AUG-2002; 2002US-0400689P.

XX 05-AUG-2003; 2003US-00634548.

XX (MONS) MONSANTO TECHNOLOGY LLC.

XX Norris SR, Lincoln K, Abad MS, Eilers R, Hartuysen KK;

XX Hirschberg J, Karunanandaa B, Moshiri F, Stein JC, Valentin HE;

XX Venkatesh TV;

XX WPI; 2004-157125/15.

XX New phytochrome kinase polynucleotides, useful in mediating tocopherol

XX biosynthesis and in producing plants with increased drought resistance.

XX Claim 1; SEQ ID NO 45; 189pp; English.

XX The invention relates to a novel substantially purified nucleic acid

XX molecule encoding a phytochrome kinase. The nucleic acid molecules and

XX polypeptides of the invention may be useful in mediating tocopherol

XX biosynthesis and in producing plants with increased drought resistance.

XX The current sequence is that of a phytochrome kinase protein of the invention.

XX Sequence 302 AA;

XX Query Match 55.8%; Score 857.5; DB 8; Length 302;

XX Best Local Similarity 60.6%; Pred. No. 1.9e-86;

XX Matches 175; Conservative 44; Mismatches 51; Indels 19; Gaps 5;

Oy 19 NNSLTHH-RF-CSPGFLISSPCFGLTGMGATOL--RARRSLISSAVATNSLHDVGA 73

Db 26 NSPTNHTVRFLCSP-----GVPPAVRLDQLRPPVPGAGA-EDLLYNAGA 71

Oy 74 TVAVLGGAYALVSPESITKRVNIOQSLRSKLVHLSGLFLVAMPFSGSTEARYPFAV 133

Db 72 TVGVLGGAYALVRAFDLITRNNILQOGLSRKLVHLSGLFLVAMPFSGSTEARYPFAV 131

Oy 134 VPLVNGSLTVNGSLISPSNMLIKSVTRREGABELKGPLFYVALALSAVFPFRESITG 193

Db 132 VPLVNGSLTVNGSLISPSNMLIKSVTRREGABELKGPLFYVALALSAVFPFRESITG 191

Oy 194 MISLMMCGGDCIADIMRKGSTKIPYNPKNAGSISMPFISFISALTYSSLSGY 253

Db 192 VISLMMCGGDCIADIMRKGSTKIPYNPKNAGSISMPFISFISALTYSSLSGY 251

Oy 254 LHMWETTLQRYA MVSNAVATVESLPTDQDDNISVPLATITLAAYISFGY 302

Db 252 VQDMWETTLQRYA MVSNAVATVESLPTDQDDNISVPLATITLAAYISFGY 300

RESULT 12

```

ADJ98208
ID ADJ98208 standard; protein; 314 AA.
XX
AC ADJ98208;
XX
DT 06-MAY-2004 (first entry)
XX
DE Rice phyto kinase protein.
XX
KW phyto kinase; tocopherol biosynthesis; plant; drought resistance;
XX enzyme; rice.
XX
OS Oryza sativa.
XX
PN MO2004013312-A2.
XX
PD 12-FEB-2004.
XX
PF 05-AUG-2003; 2003WO-US025276.
XX
PR 05-AUG-2002; 2002US-0400689P.
PR 05-AUG-2003; 2003US-00634548.
XX
PA (MONS ) MONSANTO TECHNOLOGY LLC.
XX
PI Norris SR, Lincoln K, Abad MS, Bjlers R, Hartuyker KK;
PI Hirschberg J, Karunanandaa B, Moshiri F, Stein JC, Valentin HE;
PI Venkatesh IV;
DR MPI; 2004-157125/15.
XX
PT New phyto kinase polynucleotides, useful in mediating tocopherol
XX biosynthesis and in producing plants with increased drought resistance.
XX
PS Claim 1; SEQ ID NO 46; 189pp; English.
XX
CC The invention relates to a novel substantially purified nucleic acid
XX molecule encoding a phyto kinase. The nucleic acid molecules and
XX polypeptides of the invention may be useful in mediating tocopherol
XX biosynthesis and in producing plants with increased drought resistance.
XX CC The current sequence is that of a phyto kinase protein of the invention.
XX
SQ Sequence 314 AA;
Query Match 51.5%; Score 791.5; DB 8; Length 314;
Best Local Similarity 51.1%; Pred. No. 4,7e-79; Indels 11; Gaps 4;
Matches 159; Conservative 60; Mismatches 81;
2 AATPLSPINHQLCR---FGNNSL-----TTRFCSPGFLISSPCFGLT-GMGSAOTDRA 53
3 AARVVDVVRHPPCSSSVAASSSLLSRSKSLASPAAAAASMRRLVGVGA---A 59
4 RRLSSAVANSLHDGATVAVGAYALVLSRESLTKRNVIOOSLSRKLVLHLSGL 113
5 PAVVAALASATPAARDCAATLLITAGAYSLVARPDGLRRLIIONSRKIVHLSGLV 119
6 PAVVAALASATPAARDCAATLLITAGAYSLVARPDGLRRLIIONSRKIVHLSGLV 119
114 PVLAMPISGSTEAPFAFVPLVNGLRVINGLSISPSMUKSVTREGRAEELKGPL 173
120 FMSSPPLSNSSTEARFAFVPLVNGLRVINGLSISPSMUKSVTREGRAEELKGPL 179
174 FVVLALLPSAVFPMWESPISGMTISLMMCGDGIADIMGRKFSSTIKIPYPRKSMAGSISM 233
180 YVIVLVAVLVFWMQSPIGIVISLMSGSGDFADIVGRYSATKLPFNKNSWIGSISM 239
234 FTFGGFISIALLYYSSIGYLMNMTTQORVAMVSMVATVESLPITDQLDNNTISVPLA 293
240 FISGFLSLALMLFYESCIGYFTVCWDLALGKLALVALAATVVEICPVNDVNDISVPLA 299
QY 294 TILAAVLSFGY 304
DB 300 TMLAAVLSFGY 310

```

```

RESULT 13
ADX93121
ID ADX93121 standard; protein; 332 AA.
XX
AC ADX93121;
XX
DT 21-APR-2005 (first entry)
XX
DE Plant full length insert polypeptide seqid 55785.
XX
KW plant protectant; plant growth regulator; gene therapy; plant;
KW recombinant DNA construct; physical array; plant breeding marker;
KW cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;
KW extreme osmotic condition; pathogen tolerance; pest tolerance;
KW growth rate; cell cycle pathway; disease resistance;
KW galactomannan production; lignin production; plant growth regulator;
KW yield; plant growth; plant development; seed oil; protein yield;
KW protein content.
XX
OS Unidentified.
XX
PN US2004034888-A1.
XX
PD 19-FEB-2004.
XX
PF 28-APR-2003; 2003US-00425114.
XX
PR 06-MAY-1999; 99US-00304517.
PR 05-NOV-2001; 2001US-00985678.
XX
PA (LIU/J) LIU J.
PA (ZHOU/Y) ZHOU Y.
PA (KOVA/) KOVALIC D K.
PA (SCRE/) SCREEN S B.
PA (TABBA/) TABASKA J E.
PA (CAO/Y) CAO Y.
XX
PI Liu J, Zhou Y, Kovalic DK, Screen SB, Tabaska JE, Cao Y;
DR MPI; 2004-180133/17.
XX
PT New recombinant DNA construct, useful for improving plant tolerance to
XX pest, cold, drought, herbicides, extreme osmotic conditions, pathogens or
XX pests, for conferring increased resistance to plant disease, or for
XX improving yield.
XX
PS Claim 1; SEQ ID NO 55785; 15pp; English.
XX
CC The invention describes a recombinant DNA construct comprising a
XX polynucleotide consisting of a sequence encoding an amino acid sequence
XX available in electronic form from the US patent office at
XX ftp.secdara.uspto.gov/sequence.html?docid:2004034888. The polynucleotide
XX of the invention are also useful in physical arrays of molecules and as
XX plant breeding markers. The recombinant DNA construct is useful for
XX improving plant tolerance to cold, heat, drought, herbicides, extreme
XX osmotic conditions, pathogens or pests, for manipulating growth rate in
XX plant cells by modification of the cell cycle pathway, for conferring
XX increased resistance to plant disease, for producing galactomannan,
XX lignin or plant growth regulators, for increasing the rate of homologous
XX recombination in plants, for improving yield by modification of
XX photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake
XX or by providing improved plant growth and development under at least one
XX stress condition or for modifying seed oil or protein yield and/or
XX content. This is the amino acid sequence of a plant full length insert
XX polypeptide that can be used in the recombinant DNA construct of the
XX invention.
XX
SQ Sequence 332 AA;
Query Match 51.5%; Score 791.5; DB 8; Length 332;
Best Local Similarity 51.1%; Pred. No. 5e-79; Indels 11; Gaps 4;
Matches 159; Conservative 60; Mismatches 81;

```



```

QY 2 AATLPLSPINHQLCR---FGNNSL-----TTHRFCSPGFLISSPCFIGHT-GMGSATQURA 53
D 21 AAARPVDVVRHPCSSSVAAASSSLLSRSKSLASPAALAAASSMRRLVLGVGAAA---A 77
QY 54 RRSIISSAVATNSLLHDVGATVAALVGAYALVLSFESLTKGNVIOQSLSRKLVHILSGIL 113
D 78 PAVAAALASATPALRDCAATLLITPAGAYSLVRAFDGLTARLLIQNLSRKIVHLSGVL 137
QY 114 FVLAMPFSGSTEARYPAAFPVLVNGRLVINGLSISFNSMLIKSVTEGSAEILLKQPL 173
D 138 FMSWPLFNSNSTEARFPAIVPLNLCIRLLTYGRLSTDENLVKSVTEGKPEELLRQPL 197
QY 174 FYVALLESANFPMRESPIGMSILAMCGGDIADIMGRKRGSTKIPYNPRKSNAGSISM 233
D 198 YVIVLVSVLVFWRQSPIGIVLSMMSGGDPADIVGRRYGAALPENENKSMIGSISM 257
QY 234 FIFGFIFISALLYYSSIGYLMNMETTLQRYAVMSVATVESLPITDQDDNISVPLA 293
D 258 FISGFLSALMLFYPSCLGYFTVCMDLALGKALVALAATVVECI PVNDVVDNISVPLA 317
QY 294 TILAAVLSFGY 304
D 318 TMLAAVLLFGY 328

```

RESULT 14

ADJ98211
ID ADJ98211 standard; protein; 657 AA.

```

XX AC ADJ98211;
XX DT 06-MAY-2004 (first entry)
XX DE Rice phyto1 kinase protein.
XX KW phyto1 kinase; tocopherol biosynthesis; plant; drought resistance;
XX KM enzyme; rice.
XX OS Oryza sativa.
XX PN MO2004013312-A2.
XX PD 12-FEB-2004.
XX PE 05-AUG-2003; 2003WO-US025276.
XX PR 05-AUG-2002; 2002US-0400689P.
XX PR 05-AUG-2003; 2003US-00634548.
XX PA (MONS ) MONSANTO TECHNOLOGY LLC.
XX PI Norris SR, Lincoln K, Abad MS, Eilers R, Harteauyker KK,
PI Hirschberg J, Karunanandaa B, Moshiri F, Stein JC, Valentin HB,
PI Venkatesh TV;
XX WP1; 2004-157125/15.
XX DR
XX PT New phyto1 kinase polynucleotides, useful in mediating tocopherol
XX biosynthesis and in producing plants with increased drought resistance.
XX PS Claim 1; SEQ ID NO 49; 189pp; English.
XX CC The invention relates to a novel substantially purified nucleic acid
XX molecule encoding a phyto1 kinase. The nucleic acid molecules and
XX polypeptides of the invention may be useful in mediating tocopherol
XX biosynthesis and in producing plants with increased drought resistance.
XX CC The current sequence is that of a phyto1 kinase protein of the invention.
XX SQ Sequence 657 AA;

```

Query Match 51.5%; Score 791.5; DB 8; Length 657;
Best Local Similarity 51.1%; Pred. No. 1.4e-78;
Matches 159; Conservative 60; Mismatches 81; Indels 11; Gaps 4;

```

QY 2 AATLPLSPINHQLCR---FGNNSL-----TTHRFCSPGFLISSPCFIGHT-GMGSATQURA 53
D 346 AAARPVDVVRHPCSSSVAAASSSLLSRSKSLASPAALAAASSMRRLVLGVGAAA---A 402
QY 54 RRSIISSAVATNSLLHDVGATVAALVGAYALVLSFESLTKGNVIOQSLSRKLVHILSGIL 113
D 403 PAVAAALASATPALRDCAATLLITPAGAYSLVRAFDGLTARLLIQNLSRKIVHLSGVL 462
QY 114 FVLAMPFSGSTEARYPAAFPVLVNGRLVINGLSISFNSMLIKSVTEGSAEILLKQPL 173
D 463 FMSWPLFNSNSTEARFPAIVPLNLCIRLLTYGRLSTDENLVKSVTEGKPEELLRQPL 522
QY 174 FYVALLESANFPMRESPIGMSILAMCGGDIADIMGRKRGSTKIPYNPRKSNAGSISM 233
D 523 YVIVLVSVLVFWRQSPIGIVLSMMSGGDPADIVGRRYGAALPENENKSMIGSISM 582
QY 234 FIFGFIFISALLYYSSIGYLMNMETTLQRYAVMSVATVESLPITDQDDNISVPLA 293
D 583 FISGFLSALMLFYPSCLGYFTVCMDLALGKALVALAATVVECI PVNDVVDNISVPLA 642
QY 294 TILAAVLSFGY 304
D 643 TMLAAVLLFGY 653

```

RESULT 15

ADJ98210
ID ADJ98210 standard; protein; 803 AA.

```

XX AC ADJ98210;
XX DT 06-MAY-2004 (first entry)
XX DE Rice phyto1 kinase protein.
XX KW phyto1 kinase; tocopherol biosynthesis; plant; drought resistance;
XX KM enzyme; rice.
XX OS Oryza sativa.
XX PN MO2004013312-A2.
XX PD 12-FEB-2004.
XX PE 05-AUG-2003; 2003WO-US025276.
XX PR 05-AUG-2002; 2002US-0400689P.
XX PR 05-AUG-2003; 2003US-00634548.
XX PA (MONS ) MONSANTO TECHNOLOGY LLC.
XX PI Norris SR, Lincoln K, Abad MS, Eilers R, Harteauyker KK,
PI Hirschberg J, Karunanandaa B, Moshiri F, Stein JC, Valentin HB,
PI Venkatesh TV;
XX WP1; 2004-157125/15.
XX DR
XX PT New phyto1 kinase polynucleotides, useful in mediating tocopherol
XX biosynthesis and in producing plants with increased drought resistance.
XX PS Claim 1; SEQ ID NO 48; 189pp; English.
XX CC The invention relates to a novel substantially purified nucleic acid
XX molecule encoding a phyto1 kinase. The nucleic acid molecules and
XX polypeptides of the invention may be useful in mediating tocopherol
XX biosynthesis and in producing plants with increased drought resistance.
XX CC The current sequence is that of a phyto1 kinase protein of the invention.
XX SQ Sequence 803 AA;

```

Query Match 51.5%; Score 791.5; DB 8; Length 803;
Best Local Similarity 51.1%; Pred. No. 1.8e-78;

Matches 159; Conservative 60; Mismatches 81; Indels 11; Gaps 4;

```
Qy 2 AATPLPSINQLCR---FGNNSL----TTHRFCSPGFLISSPCFGLT-GMGSATQLRA 53
Db 492 AARPVDVVRHPCCSSVVAASSSLLSRSKSLASPAALAAASMRRLVLGVGAAA--A 548
Qy 54 RRSLLSSAVATNSLHDVGVATVAVLGAYALVLSFESLTKRNVIQOSLSRKLVIHLSGL 113
Db 549 PAVAAALASATPPALRDCATLLITAGAYSIVRAFDGLTARRLIEQNLSRKIVHLSGVL 608
Qy 114 FVLAMPFSSGSTEARFPAFVPLVNGLRVLVINGLSISPNMILKSVTREGRAEELKGPL 173
Db 609 FMSWPLFSNSTEARFFPAIIVPLNCIRLLTYGLRLSTDEALVKSVTREGKPEEELRGPL 668
Qy 174 FVVLALLFSAVPFMRRESPIGMIISLAMCGGDGIADIMGRKFGSTKIPLYNPRKSNAGSISM 233
Db 669 YVIVLVLSVLPFWKQSPIGIVLSWMSGGDGFADIVGRRYGSAKLPFNENKSMIGSISM 728
Qy 234 FIFGFFISIALLYYSSLSGYLHMNMTLQORVAVSMVATVSESLLPTDOLDNISVPLA 293
Db 729 FISGFLLSALMLFYSCIGYFTVCMDLALGKLALVALAATVVECI PVNDVVDNISVPLA 788
Qy 294 TILAAYLSFGY 304
Db 789 TMLAAYLLFGY 799
```

Search completed: March 16, 2006, 19:49:02
Job time : 92 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 16, 2006, 15:34:32 ; Search time 7827 Seconds
(without alignments)
7923.369 Million cell updates/sec

Title: US-10-634-548-1

Perfect score: 1091
Sequence: 1 aaaaaagaataataaccacaa.....ccgaacttaaatgttcc 1091

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl: *
1: gb_ba: *
2: gb_in: *
3: gb_env: *
4: gb_om: *
5: gb_ov: *
6: gb_pat: *
7: gb_ph: *
8: gb_pr: *
9: gb_to: *
10: gb_str: *
11: gb_gy: *
12: gb_un: *
13: gb_vl: *
14: gb_hlg: *
15: gb_pl: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1075	98.5	1091	15	AY085036 Arabidops
2	1051.4	96.4	1072	15	BT004006 Arabidops
3	1002	91.8	1004	15	BT021123 Arabidops
4	404	37.0	93045	15	AT132M21 Arabidops
5	161.4	14.8	1002	15	AK061265 Oryza sat
6	147	13.5	1218	15	AK10748 Oryza sat
7	144.8	13.3	1059	15	AK109049 Oryza sat
8	138	12.6	1211	15	AK176090 Arabidops
9	136.4	12.5	1360	15	AK176217 Arabidops
10	134.8	12.4	1162	15	AY087555 Arabidops
11	127	11.6	436	15	AK063069 Oryza sat
12	86	7.2	110000	14	CR954207 Ostrreococ
13	79	6.2	81672	15	AB020755 Arabidops
14	67.2	6.1	7218	6	166494 Sequence 14
15	67	6.0	3143	14	OSIG00051 Arabidops
16	65.8	6.0	95419	14	OSIG00051 Oryza sat
17	65.8	6.0	110000	15	AP008210_342 Continuation (343

19	65.8	6.0	190432	15	OSJN00032	AL606619 Oryza sat
20	58.8	5.4	477	10	BV151380	BV151380 P2A02133-
21	58.8	5.4	465	10	BV151381	BV151381 P2A02133-
22	58.8	5.4	493	10	BV151387	BV151387 P2A02133-
23	58.8	5.4	506	10	BV151376	BV151376 P2A02133-
24	58.8	5.4	506	10	BV151384	BV151384 P2A02133-
25	58.8	5.4	513	10	BV151379	BV151379 P2A02133-
26	58.8	5.4	513	10	BV151382	BV151382 P2A02133-
27	58.8	5.4	515	10	BV151385	BV151385 P2A02133-
28	58.8	5.4	516	10	BV151377	BV151377 P2A02133-
29	58.8	5.4	518	10	BV151383	BV151383 P2A02133-
30	58.8	5.4	529	10	BV151386	BV151386 P2A02133-
31	57.2	5.2	513	10	BV151378	BV151378 P2A02133-
32	46.8	4.3	14603	2	AC116989	AC116989 Dictyoste
33	46	4.2	200983	9	AC164158	AC164158 Mus muscu
34	45.8	4.2	129719	8	AL358113	AL358113 Human DNA
35	45.4	4.2	81410	8	AC073594	AC073594 Homo sapi
36	45.2	4.1	93331	8	AC107392	AC107392 Homo sapi
37	45.2	4.1	188941	14	AC129362	AC129362 Rattus no
38	45	4.1	165576	9	AC134587	AC134587 Mus muscu
39	45	4.1	176044	9	AC131687	AC131687 Mus muscu
40	45	4.1	253240	14	AC112378	AC112378 Rattus no
41	44.6	4.1	110000	15	AP008218	Continuation (3 of
42	44.6	4.1	152222	15	CNS08CDK	BX000494 Oryza sat
43	44.2	4.1	245390	9	AC103947	AC103947 Mus muscu
44	44	4.0	1907	2	AB089904	AB089904 Dictyoste
45	43.8	4.0	185029	14	AC068986	AC068986 Homo sapi

ALIGNMENTS

RESULT 1	AY085036	1091 bp	mRNA	linear	PLN 14-APR-2003
LOCUS	Arabidopsis thaliana clone 125255	mrna	complete sequence.		
DEFINITION	AY085036				
ACCESSION	AY085036.1	GI:21403746			
VERSION					
KEYWORDS	FLI CDNA.				
SOURCE	Arabidopsis thaliana (thale cress)				
ORGANISM	Arabidopsis thaliana				
REFERENCE	Haas,B.J., Volkovskiy,N., Town,C.D., Troughan,M., Alexandrov,N., Feldmann,K.A., Flavell,R.B., White,O. and Salzberg,S.L.				
AUTHORS	Full-length messenger RNA sequences greatly improve genome annotation				
TITLE	Genome Biol. 3 (6), RESEARCH0029 (2002)				
JOURNAL	12093376				
PUBMED	2 (bases 1 to 1091)				
REFERENCE	Brover,V., Troughan,M., Alexandrov,N., Lu,Y.-P., Flavell,R. and Feldmann,K.				
AUTHORS	Full-length cDNA from Arabidopsis thaliana				
TITLE	Unpublished				
JOURNAL	3 (bases 1 to 1091)				
REFERENCE	Brover,V., Troughan,M., Alexandrov,N., Lu,Y.-P., Flavell,R. and Feldmann,K.				
AUTHORS	Direct Submission				
TITLE	Submitted (11-MAR-2002) Ceres, Inc, 3007 Malibu Canyon Road,				
JOURNAL	Malibu, CA 90265, USA				
COMMENT	This clone sequence is one of 5,000 Ceres full-length cDNAs made available to TIGR and Genbank. The following quality assessment of this set was done by comparison with known proteins: two percent of the clones are estimated to be 5'-truncated; less than one percent are 3'-truncated; approximately two percent represent alternative splice variants; including unspliced introns and spliced exons; one percent may contain premature stop codons; five percent may have frame shifts in a coding region. A sequence is considered to be 5'-truncated if it lacks the translation initiation start (ATG). A sequence is considered to be 3'-truncated if it lacks the C-terminal end of the encoded protein. Please note that these cDNA				

960 GGGTCACTAGATACCTTTACAACCAGAAGAAAGATTGSGGAGGAAGATCTCCATGTTC 809
961 GGGTCACTAGATACCTTTACAACCAGAAGAAAGATTGSGGAGGAAGATCTCCATGTTC 720
962 ATCTTCGGCTTCCTCATCTGCATCGCATTAATTACTAATTACTCAAGCCTTGSGTACTT 869
963 ACTTCGGCTTCCTCATCTGCATCGCATTAATTACTAATTACTCAAGCCTTGSGTACTT 780
964 CACATGAACCTGGGAAAACGACCTTGACAGAGTAGCAATATTTGGCTTCGCCACGGTA 929
965 CACATGAACCTGGGAAAACGACCTTGACAGAGTAGCAATATTTGGCTTCGCCACGGTA 840
966 GTCCAGTGGCTCACCCATCACCGATCAATTAGACGACATATTTGGCTTCGCCACTACT 900
967 GTCCAGTGGCTCACCCATCACCGATCAATTAGACGACATATTTGGCTTCGCCACTACT 900
968 ATTATGAGCTGCTTATTTAAGTTTCGAGATATTTAGATTAATTCCTCATAAACGAAATGTGA 1049
969 ATTATGAGCTGCTTATTTAAGTTTCGAGATATTTAGATTAATTCCTCATAAACGAAATGTGA 1050
970 TATAGCTATTTTTTAAATGAATCCGACCTTCAAAATGTTTCC 1091
971 TATAGCTATTTTTTAAATGAATCCGACCTTCAAAATGTTTCC 1002
972 TATAGCTATTTTTTAAATGAATCCGACCTTCAAAATGTTTCC 1002

ATTJ32M21 93045 bp DNA linear PLN 16-APR-2005
Arabidopsis thaliana DNA chromosome 5, BAC clone TJ32M21 (ESSA project).
AL162875
AL162875.1 GI:7406444

Arabidopsis thaliana (thale cress)
Arabidopsis thaliana
Bakerfycia; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophytes; Magnoliophyta; eudicotyledons; core eudicotyledons;
Rosids; eucosids II; Brassicales; Brassicaceae; Arabidopsi.
1 (bases 1 to 93045)
Bevan,M., Terryn,N., Ardiles,W., Buysehaert,C., Dasseville,R., De
Clerck,R., De Keyser,A., Neyt,P., Rouze,P., Van Den Daele,H.,
Villareel,R., Gielens,J., Van Montagu,M., Bancroft,I., Mewes,H.W.,
Rudd,S., Lemcke,K. and Mayer,K.F.X.
Unpublished
2 (bases 1 to 93045)
EU Arabidopsis sequencing project.
Direct Submission
Submitted (31-MAR-2000) MIPS, at the Max-Planck-Institut fuer
Biochemie, Am Klopferspitz 18a, D-82152 Martinsried, FRG, E-mail:
lemcke@mip.biochem.mpg.de, mayev@mip.biochem.mpg.de Project
Coordinator: Mike Bevan, Molecular Genetics Department, Cambridge
Laboratory, John Innes Centre, Colney Lane, NR4 7UJ Norwich, UK,
E-mail: michael.bevan@bsrc.ac.uk
Information on performance of analysis and a more detailed
annotation of this entry and other sequences of chromosomes 3, 4
and 5 can be viewed at: http://www.mips.biochem.mpg.de/proj/thal/.
Location/Qualifiers
1..93045
/organism="Arabidopsis thaliana"
/mol_type="Genomic DNA"
/db_xref="taxon:3702"
/chromosome="5"
/ecotype="Columbia"
1..1000
/note="this clone overlaps with BAC TJ9M18"
676..3112
/gene="TJ32M21_10"
join(676..841,958..1235,1354..1545,1884..2447,2525..2642,
2727..3112)
/gene="TJ32M21_10"
/note="strong similarity to NAC2, Arabidopsis thaliana,
EMBL:AF201456"

```

/codon_start=1
/product="putative protein"
/protein_id="CAB8547.1"
/db_xref="GI:7406445"
/db_xref="GOA:Q84K00"
/db_xref="InterPro:IPR003441"
/db_xref="UniProt/Swiss-Prot:Q84K00"
/translation="MGRGSVTSILAGFRFHPDDELRYYLKRVCKNPKPKDAISVT
DIVKSPMDLPDQSKLKRDLKRWYPMRLDKYNGSKTRATKGVKTKTGDREIR
NGSRVVGKKTLLVTHKRAPRGRITNMVHEIRLSDDLKAGVQEAIVLCRIQKRS
GTGPKNGEQGAPYLEBWEEDGMTYVPAQAFSEGLANDVYVDIDIDEKENLY
VYDAVPLPNYCHGESSNNVSGNYSQSGNTIQGNVNVDSGGYEQPIRTFEDRFP
IIRGSGIOPCSLPEBOIGCGVDENVLSSNNNYFVADTCSDIPIDHNVLPDPP
FMDPNMNLPLNDGLYLETNDLSCAQODDFNEDYLSFDDGLFPDLSILMGPEDFLP
NOGLDOKRAPKLEKEVAGGKAEVKEGSGSSKQDDDFRQDPSAPKYPRLKTS
HMLGALPTPSSFASQFQTKDMRLHAAQSSGSVHTVIGMMKISMTTLAUSGMSYTD
KNGMLNVLSFGVQDDAMTASGSKTGITATRAMLVFMCILVLLSVSRKIVTWASA
R"
/exon
676..841
/gene="TJ32M21_10"
/number=1
642..957
/gene="TJ32M21_10"
/number=1
958..1235
/gene="TJ32M21_10"
/number=2
1236..1353
/gene="TJ32M21_10"
/number=2
1354..1545
/gene="TJ32M21_10"
/number=3
1546..1883
/gene="TJ32M21_10"
/number=3
1884..2447
/gene="TJ32M21_10"
/number=4
2448..2524
/gene="TJ32M21_10"
/number=4
2525..2642
/gene="TJ32M21_10"
/number=5
2643..2726
/gene="TJ32M21_10"
/number=5
2727..3112
/gene="TJ32M21_10"
/number=6
3113..6151
/gene="TJ32M21_20"
complement(join(3563..3721,3811..3870,3949..4269,
4349..4432,4513..4601,4709..4823,4900..5026,5142..5281,
5368..5429,5507..5573,5652..5783,5870..5947,6041..6151))
/gene="TJ32M21_20"
complement(join(3563..3721,3811..3870,3949..4269,
4349..4432,4513..4601,4709..4823,4900..5026,5142..5281,
5368..5429,5507..5573,5652..5783,5870..5947,6041..6151))
/notes="similarity to rngB protein, Dictyostelium
discoideum, PIR:S68824"
/codon_start=1
/product="putative protein"
/protein_id="CAB8548.1"
/db_xref="GI:7406446"
/db_xref="InterPro:IPR006652"
/db_xref="UniProt/TREMBL:Q9L283"
/translation="MEIRKREIDVGDHNSNLADDEWTPLPVSGSRASARYKHAAYVD
EKYIYVGSRRGRALSDVOVFDLRLSLTWSSIKLKTSSSADNIOEDDGSIRKAFPAI
SDHRMIMGNKLTLLIGHSKSSSDNMLVRFIDLETSHSCGVLDVFGNVPASRGHSITL

```

```

/exon
complement(3563..3721)
/gene="TJ32M21_20"
/number=1
complement(3722..3810)
/gene="TJ32M21_20"
/number=1
complement(3811..3870)
/gene="TJ32M21_20"
/number=2
complement(3871..3948)
/gene="TJ32M21_20"
/number=2
complement(3949..4269)
/gene="TJ32M21_20"
/number=3
complement(4270..4348)
/gene="TJ32M21_20"
/number=3
complement(4349..4432)
/gene="TJ32M21_20"
/number=4
complement(4433..4512)
/gene="TJ32M21_20"
/number=4
complement(4513..4601)
/gene="TJ32M21_20"
/number=5
complement(4602..4708)
/gene="TJ32M21_20"
/number=5
complement(4709..4823)
/gene="TJ32M21_20"
/number=6
complement(4824..4899)
/gene="TJ32M21_20"
/number=6
complement(4900..5026)
/gene="TJ32M21_20"
/number=7
complement(5027..5141)
/gene="TJ32M21_20"
/number=7
complement(5142..5281)
/gene="TJ32M21_20"
/number=8
complement(5282..5367)
/gene="TJ32M21_20"
/number=8
complement(5368..5429)
/gene="TJ32M21_20"
/number=9
complement(5430..5506)
/gene="TJ32M21_20"
/number=9
complement(5507..5573)
/gene="TJ32M21_20"
/number=10
complement(5574..5651)
/gene="TJ32M21_20"
/number=10
complement(5652..5783)
/gene="TJ32M21_20"
/number=11
complement(5784..5869)
/gene="TJ32M21_20"
/number=11

```


QY 306 CTGTGATGACGTCGAGACCGGTGCACTGCTGGGAGATACGGCTTGTCTA 365
 DB 186 CTGGCCCAACGCTCGGCTCGCGCCCTTACCGGAGCGCTCGCCCTCTCCGCC 245
 QY 366 AGCTTCAGAGTCTCACCAGCAAGAGTCAATTCAGAGTTCAGAGCAAGAGTTCG 425
 DB 246 TTCTTCAGAGAGCTCGCAGAGCGGGGGTCTTCAGAGCAAGAGTCAATTCAGAG 305
 QY 426 CATTAATCTCAAGTCTGCTTTTGTGACTTCGCGCAATCTTCAGCGATTCAGCGAG 485
 DB 306 CATTAATCTCAAGTCTGCTTTTGTGACTTCGCGCAATCTTCAGCGATTCAGCGAG 365
 QY 486 GCTGATACCTTGTGCTTTTGTGACTTCGCGCAATTCAGCGATTCAGCGAG 545
 DB 366 GCTGATACCTTGTGCTTTTGTGACTTCGCGCAATTCAGCGATTCAGCGAG 425
 QY 546 CTATCAATTCCTCCCAATTCGATGCTAATTCCTGCAAGAGAGAGAGAGAG 605
 DB 426 TTAGAGAGTAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 485
 QY 606 GAGTTGCTTAAAGTCTTTTGTGACTTCAGCTCTTTTCTCTGCGGTTTCTTC 665
 DB 486 GAACTTCTCAAGGCTCCATGATATGCTACCACTAATCTTTCGCACTTCTATATTC 545
 QY 666 TGGAGAGAGTCTCTATCGGATGATCTCGTACCAATGATGCTGCGGATGAGATA 725
 DB 546 TGGAGAGAGTCTCTATCGGATGATCTCGTACCAATGATGCTGCGGATGAGATA 605
 QY 726 GCTGATTAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 785
 DB 606 GCGAGACATGAG 665
 QY 786 TGGGAG 845
 DB 666 TATGAG 725
 QY 846 TATTAATCAAGGCTTGGGATCTTCAATGAGAGAGAGAGAGAGAGAGAGAG 905
 DB 726 TACTTCAAGTATTCGCTTCTTATGAGAGAGAGAGAGAGAGAGAGAGAGAG 782
 QY 906 ATGCTCAATGAG 965
 DB 783 GTTCTTCTGTAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 842
 QY 966 AATATTCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
 DB 843 AACTGACGCTTCTTAAATCAATCTTCTGCTGCTGCTGCTGCTGCTGCTGCT 897
 RESULT 6
 AK10748 1218 bp mRNA linear PLN 24-JUL-2003
 LOCUS Oryza sativa (japonica cultivar-group) cDNA clone:002-170-604, full
 DEFINITION insert sequence.
 ACCESSION AK10748
 VERSION AK10748.1 GI:32995957
 KEYWORDS Full_CDNA; oligo capping.
 SOURCE Oryza sativa (japonica cultivar-group)
 ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophytes; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Eriatoidae; Oryzaceae; Oryza.
 REFERENCE
 1
 The Rice Full-Length cDNA Consortium, National Institute of
 Agrobiological Sciences Rice Full-Length cDNA Project Team:
 Kikuchi, S., Sato, K., Nagata, T., Kawagashira, N., Doi, K.,
 Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotte, I.,
 Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C.,
 Onsenki, K., Shishiki, T., Foundation of Advancement of International
 Science Genome Sequencing & Analysis Group, Oono, Y., Murakami, K.,
 Iida, Y., Sugano, S., Fujimura, T., Suzuki, Y., Tsunoda, Y.,
 Kurosaki, T., Kodama, T., Masuda, H., Kobayashi, M., Xie, Q., Lu, M.,

TITLE
 JOURNAL
 PUBMED
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 PUBMED
 REFERENCE
 AUTHORS
 COMMENT
 URL: //cdna01.dna.affrc.go.jp/cdna/
 NIAS Rice Full-Length cDNA Project Team: Kikuchi, S., Sato, K.,
 Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J.,
 Ishikawa, M., Yamada, H., Ooka, H., Hotte, I., Kojima, K., Namiki, T.,
 Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Onsenki, K., Shishiki, T.,
 Yamamoto, M.
 PALS Genome Sequencing & Analysis Group: Oono, Y., Iida, Y.,
 Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M., Kobayashi, M.,
 Kodama, T., Kurosaki, T., Kusumegi, T., Lu, M., Masuda, H., Mura, J.,
 Mizuno, K., Narikawa, R., Nishikawa, J., Oka, M., Ryu, R., Sugano, S.,
 Sugiyama, A., Suzuki, Y., Tsunoda, Y., Ueda, M., Xie, Q., Yokomizo, S.,
 Yoshimura, A., Matsubara, K., and Murakami, K.
 Genome Exploration Research Group in Riken Genomic Sciences Center
 and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K.,
 Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hangaki, T.,
 Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K.,
 Hirooka, T., Hori, F., Iida, J., Imamura, K., Imotani, K., Ishii, Y.,
 Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawai, J.,
 Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
 Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
 Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Oono, M., Ogino, N.,
 Oka, Y., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H.,
 Sasaki, D., Sato, K., Shibata, K., Shingawa, A., Shiraki, T.,
 Sogabe, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F.,
 Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Waki, K.,
 Yasunishi, A., and Hayashizaki, Y.
 Location/Qualifiers
 1..1218
 /organism="Oryza sativa (japonica cultivar-group)"
 /mol_type="mRNA"
 /cultiyar="Nipponbare"

Db	267	TCCTGCTGCTGGGTTTCTGGGGAGAGATTGGCAACGTCGATCTTGCACCGAATCTCA	326
QY	412	GCAGAAAGCTTGTGCATTAATCTTCAGGCTCTGCTTTTGTACTTGGCTGGCCAAATCTTCA	471
Db	327	TCCGGAAGCTTGTGCATTAATAATATTGGGCTAGTTTATTATGCTTTTGTGGCCGCTGTCA	386
QY	472	GCGGATGCACCGAGGCTCGATTACTTGTGCTTTTGTTCCTGGTATGTAATGGCTTAAAGC	531
Db	387	GTTCTGGAATCCAAAGAGACACTTTCCTGATCTCTTGTACTCGGAATCAATATAGTAAGA	446
QY	532	TTGTATTAATACGAGACTATTCATTTTCCCAAAATTGCATGCTATCAATCCGTCACAAAG	591
Db	447	TGCTATTCCTGGGGCTTGGAGTGTACACGACGAAGAAACATCAAGTCAATGAGCAGAC	506
QY	592	AAGGAGAGACGAAGAAGTTGCTTAAAGGTCCTTTGTTTACGTTTACCTCTTCTTTCT	651
Db	507	ATGGAGATCGACGGGAATCTATTAAAGGGGCGCTTACTATAGTATGTCAATCATCAG	566
QY	652	CTGCGGTTTTCTTCTGAGAGAGTCTCTTAATCGGATGATCTCGTTAGCAATGATGTGT	711
Db	567	CTGTGCTACTATTTGGAATATCATCTCCAAATCGCAATGCGGTGATATGCAACCTTTGCG	626
QY	712	GTTGGGATGGAATATGCTATATTAATAGGAGCGTAAGTTGGGTCACATAAGATACCTTACA	771
Db	627	CAGGAGATGTATGCTCAATGTGGGTGCGCGGTTTGGAAACAGAGAAGCTTCTTACCA	686
QY	772	ACCCAAGAAAGTTGGGGCAGAGACATCTCCATTTGCTATCTTGGCTTCTTCAATCTTCA	831
Db	687	ACAAABAAACAAATCTTCTGCTGTAGCATTTGAAATGGCCACCGCGGTTTCTAGATCTG	746
QY	832	TCGCATTACTTACTTAATCTCAAGCCTTGGGTACCTTTCACATGAACTGGGAAACGACT	891
Db	747	TTGCGCTATATGTACTACTTGTCTTCAATTTGGTATATGAGAGATAGCGGGGAATGA---	803
QY	892	TGCAAGAGTAGCATATGCTCTCAATGCTGCGCAAGTATGAGATGCTAACCCATCAACG	951
Db	804	TTCTTGTTTCTCGTCAATCTTATAGATCAAGCTCTTGTGAAATCACTCTCCAAATAGCA	863
QY	952	ATCAATTAGACGACAAATATTTTCGTTCCCTGCTGCTACTAATTTTATAGCTCTTATTAAGTT	1011
Db	864	CCGACATTGACGACAAATCTCAACATTTCTTAACTCTGCTTGGCCGGATTTCTTACTCT	923
QY	1012	TCGATATTA 1021	
Db	924	TCTAATAATA 933	
RESULT 9			
AKI176217		1360 bp mRNA linear	PLN 09-SEP-2004
LOCUS		Arabidopsis thaliana mRNA, complete cds, clone: RAFL23-09-006.	
DEFINITION			
ACCESSION		AKI176217	
VERSION		AKI176217.1 GI:51970575	
KEYWORDS		FLI CDNA.	
SOURCE		Arabidopsis thaliana (thale cress)	
ORGANISM		Arabidopsis thaliana	
		Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;	
		Spermatophytes; Magnoliophyta; eudicotyledons; core eudicotyledons;	
		rosoids; euroside II; Brassicales; Brassicaceae; Arabidopsis.	
REFERENCE		1	
AUTHORS		Tsuchi, Y., Seki, M., Ishida, J., Nakajima, M., Enju, A., Kamiya, A.,	
		Narusaka, M., Shin-i, T., Nakagawa, M., Sakamoto, N., Oiehi, K.,	
		Kohara, Y., Kobayashi, M., Toyoda, A., Sakaki, Y., Sakurai, T., Iida, K.,	
		Akiyama, K., Satou, M., Toyoda, T., Konagaya, A., Carninci, P.,	
		Kawai, J., Hayashizaki, Y. and Shinozaki, K.	
TITLE		Large-scale analysis of RIKEN Arabidopsis full-length (RAFL) cDNAs	
JOURNAL		Unpublished	
REFERENCE		2 (bases 1 to 1360)	
AUTHORS		Tsuchi, Y., Seki, M., Ishida, J., Nakajima, M., Enju, A., Kamiya, A.,	
		Narusaka, M., Shin-i, T., Nakagawa, M., Sakamoto, N., Oiehi, K.,	
		Kohara, Y., Kobayashi, M., Toyoda, A., Sakaki, Y., Sakurai, T., Iida, K.,	
		Akiyama, K., Satou, M., Toyoda, T., Konagaya, A., Carninci, P.,	

FEATURES	source	location/Qualifiers
FEATURES	source	1..1360
		/organism="Arabidopsis thaliana"
		/mol_type="mRNA"
		/db_xref="taxon:3702"
		/chromosome="5"
		/clone="RAF123-09-006"
		/ecotype="Columbia"
		/note="common name: thale cress"
gene		1..1360
		/gene="At5g58560"
		120..1043
		/gene="At5g58560"
		/codon_start=1
		/product="hunkun protein"
		/protein_id="BAD33980.1"
		/db_xref="GI:51970576"
		/translation="MATSTSTTKLSVLCSPFISPLVDSPSLAFSPFIRELTURIA TSFSSRPATKIRKSLAAVMPENSVSDVCAFEVTSIVFSCGFNBEIGKROI FDQILKLVHINIGLVCMPLFSGIGALFASLPGLIVRMILGIVYHDGEG TISMSHGRLRINKGSPLYLVLSITGACIYKMSSPILATAVLCNLCAGMDIVR RFREKIPYNNKSPRAGISGMATNGPLASVAVMYTPRSPRIVEDSGMILRPLVISIA SALVESLPISIDIDNLTLSLALGFLP"
ORIGIN		
Query Match	12.5%; Score 136.4; DB 15; Length 1360;	
Best Local Similarity	50.1%; Pred. No. 2.5e-26;	
Matches	366; Conservative 0; Mismatches 361; Indels 3; Gaps 1;	
Db	292 TTGGACGAATTCGCTGTTCATGACGTCGAGCCACCTGGACGAGCTTGGTGAACAT	351
	322 TTCGGAAATATGGTTTATCAGATGTCGCGGTTTGGAGTCATAGCATCGTGGGT	381
Db	352 ACGGCTTGTCTTAAGCTTGAAGCTCACCAAGCAACGTATTCACAGAGTTGA	411
Db	382 TCTGTCCTCGCTTTCGGGGAGAGATGGCAACGTGGCATCTTCGACCAAGAACTCA	441
Db	412 GCAGAAAGCTTGTGCATATATCTCCAGGTCGCTTTCGTAAGTTCGCGGCAATCTTCA	471
Db	442 TCCGAAAGCTTGTGCATATATATTTGGGCTAGTTTATGCTTTCGCGCGCTTTCA	501
Db	472 GCGATGACACCGAGGCTCGATCTTTCGCTTTCGTTTGAAGATGGCTTAAGGC	531
Db	502 GTTTCGAATTCAGAGACACTTTTGCATCTCTTGAACCTGACATCATATATTAAGGA	561
Db	532 TTGTTATTAACGACATATCCATTTCCCAATTCGATGCTATCAAAATCCGTCACAGAG	591
Db	562 TGCTATTGCTGGGCTTGAAGTGAACGACGAGAAACATCAAGTCATGAGCAGAC	621
Db	592 AAGGAGAGCAGAGAGTGGTTAAAGTCTTGTCTTACGTTCTAGCTCTTCTTTCT	651
Db	622 ATGAGATTCGAGGAACTACTTAAGGCGCGCTTATCTATGATCTGCAATCAATCAG	681
Db	652 CTGGCGTTTCTTTCGAGAGAGCTCTTATCGGTATGATCTCGTTAGCAATGATGTGTG	711
Db	682 CTCGATCTACTATTGGAATCATCCCAATCGGATTTGGGTATATGACAACTTTGGG	741
Db	712 GTGCGATGAGATGCTGATATTAATGGAGCGTAATTTGGGTCAATTAAGATCACTTACA	771

```

Db      742 CAGAGATGATGATGCTACATTTGGGGTCCGGCGTTTGGAAACAGAGAGCTTCTTACA 801
Qy      772 ACCAAGAAAGATTGGGAGAGAGATCTCCATGTTTCATTTGGCTTTCTTCA 831
Db      802 ACAAACAAATCATTTGCTGGTAGCATGGAAATGGCCAGCGGGTTTCTACATCTG 861
Qy      832 TCGATTAATTACTTACTTACCAAGCCTTGGGTACCTTCAATGAACCTGGGAAGCACT 891
Db      862 TTGGGTATATGATCTACTTCTTCATTTGTTACATGAGATAGCGGGGAATGA--- 918
Qy      892 TCGAGAGATGACATGCTCTCAATGATCGCCAGAGTAGTGAAGTCCGATCCATCCG 951
Db      919 TTTCTGTTTCTGCTGATCTCTATAGATCAGCTCTGTGGATCACTCCCAATAGCA 978
Qy      952 ATCAATTAGACCAATTTTGGTCTCTGCTACTATTATTAAGTCTTAAATT 1011
Db      979 CCGACATTTAGCAACATCTCACATTTCTTAACCTTGCCCTGGCCGATTTCTACTCT 1038
Qy      1012 TCGGATATTA 1021
Db      1039 TCTAATATA 1048

```

RESULT 10
AY087555
LOCUS AY087555 1162 bp mRNA linear PLN 14-APR-2003
DEFINITION Arabidopsis thaliana clone 36602 mRNA, complete sequence.
ACCESSION AY087555
VERSION AY087555.1 GI:21406293
KEYWORDS FLI CDNA.
SOURCE Arabidopsis thaliana (chale crese)

ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; euroside II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 1162)
Haas,B.J., Volkovskiy,N., Town,C.D., Troughan,M., Alexandrov,N.,
Feldmann,K.A., Flavell,R.B., White,O. and Salzberg,S.L.
Full-length messenger RNA sequences greatly improve genome
annotation

REFERENCE
AUTHORS
TITLE
JOURNAL
PUBMED
12093376
2 (bases 1 to 1162)
3 (bases 1 to 1162)
Brover,V., Troughan,M., Alexandrov,N., Lu,Y.-P., Flavell,R. and
Feldmann,K.
Full-length cDNA from Arabidopsis thaliana
Unpublished
3 (bases 1 to 1162)
Brover,V., Troughan,M., Alexandrov,N., Lu,Y.-P., Flavell,R. and
Feldmann,K.
Direct Submission
Submitted (11-MAR-2002) Ceres, Inc, 3007 Malibu Canyon Road,
Malibu, CA 90265, USA

COMMENT
TITLE
JOURNAL
REFERENCE
AUTHORS
This clone sequence is one of 5,000 Ceres full-length cDNAs made
available to Titer and Genbank. The following quality assessment of
this set was done by comparison with known proteins: two percent of
the clones are estimated to be 5'-truncated; less than one percent
are 3'-truncated; approximately two percent represent alternative
splice variants, including unspliced introns and spliced exons; one
percent may contain premature stop codons; five percent may have
frame shifts in a coding region. A sequence is considered to be
5'-truncated if it lacks the translation initiation start (ATG). A
sequence is considered to be 3'-truncated if it lacks the
C-terminal end of the encoded protein. Please note that these cDNA
sequences are derived from the Ws or later ecotypes and therefore
may contain polymorphisms when compared to sequences from Col-0.
Genet carried out the library production and sequencing of the
full-length clones. Ceres, Inc. carried out the clustering of the
5' sequences, selection of clones, and sequence assembly.
Location/Qualifiers
1. .1162
/organism="Arabidopsis thaliana"

FEATURES
Source

```

/mol_type="mRNA"
/db_xref="taxon:3702"
/clone="36602"
23..946
/codon_start=1
/product="unknown"
/protein_id="PAM65097.1"
/db_xref="GI:21593148"
/translation="MATTSTTTKLSVLCGPFISPLVDSPPSLAFSPRIREFLVRIA
TSFSSSRFPATKIRKSLAAMPSPSVSLDVAEFTSIIVAFSCIGFWEIGRGI
FDQJLRKLVHINIGLVFVLCWPLFSSGIGALFASLIVPLNTVRMLLIGVYHDS
TIKSMRHGRRLKGPLLYVLSITSCACIYKWSPIALAVICNLGAGGMDIVGR
RFGREKLPYNNKRSFASIGMATGFAVGYMYFASFGYIEDSGMILRFVLSIA
SALVESLPISITDIDNLTISLISLAAGFLF"

```

Query Match 12.4%; Score 134.8; DB 15; Length 1162;
Best Local Similarity 50.0%; Pred. No. 7e-26;
Matches 365; Conservative 0; Mismatches 362; Indels 3; Gaps 1;

```

Qy      292 TTGGAGAAATTCGCTGTGATGACGTGGAGCCACCGTGCAGTCTGTGTGAGAT 351
Db      225 TTCCGAAAATTCGCTTTTACAGATGTGCGGCTTGAAGTCAATGATCGTGGT 284
Qy      352 ACGCGCTTGCTTAAGCTTCAGAGTCTCACCAAGCAAGTCAATCAAGAGTTGA 411
Db      285 TCTGTGCTCGCTGTTTCTGGGAGAGATTTGGCAAACTGGCATTTGACACGAACTCA 344
Qy      412 GCAGAAAGCTTGTGATATCTCTCAGATCTGCTTTTGTGATTTGCGTGGCCAACTTCA 471
Db      345 TCCGAAAGCTTGTGATATTAATATGAGGCTAGTTTATGCTTTGCTGGCCGCTTCA 404
Qy      472 GCGATGACCGAGGCTCATCTTGTGCTTTTGTTCGCTTGAATAGCTTAAGC 531
Db      405 GTTCTGAATTCAGAGGACCTTTTCCATCTCTTGAACCTGACTCAATATAGTGA 464
Qy      532 TTGTATTAACGACATATCCATTTCCCAATTCGATGATCAATCAATCCGTCACAAG 591
Db      465 TGCTATTGCTGGGCTTGAAGTATACACGAGAAACATATAGTCAATGAGCAGC 524
Qy      592 AAGGAGAGCAGAAAGTGTGTTAAAGTCTTTTGTCTAGTCTTACTCTTTTCT 651
Db      525 ATGAGATTCGAGGAACTACTTAAGGGGCGCTTACTATGATGATGATCAATCAG 584
Qy      652 CTGCGCTTTCTTGTGAGAGAGTCTCTATCGGTATGATCTGTTAGCAATATGTTG 711
Db      585 CTGCACTCTATTTGAAATCATCCCGATTCGGAATGGGATGATGCAACCTTGGC 644
Qy      712 GTGCGATGGAATAGCTGATATATAGGAGCTAAGTTGGGTCAACTAAGATACCTTACA 771
Db      645 CAGAGATGATGATGCTGACATTTGGGTGGCGGTTTGAACAGAGAGCTTCTTACA 704
Qy      772 ACCAAGAAAGATTGGGAGAGAGATCTCCATGTTTCATCTTGGCTTTCTTCA 831
Db      705 ACAAACAAATCATTTGCTGGTAGCATGGAAATGGCCAGCGGGTTTCTACATCTG 764
Qy      832 TCGATTAATTACTTACTTACCAAGCCTTGGGTACTTCAATGAACCTGGGAAGCACT 891
Db      765 TTGGGTATATGATCTACTTCTTCATTTGGTTAATCATCGAGATAGCGGGGAATGA--- 821
Qy      892 TCGAGAGATGACATGCTCTCAATGATCGCCAGAGTAGTGAAGTCCGATCCATCCG 951
Db      822 TTTCTGTTTCTGCTGATCTCTATAGATCAGCTCTTGGATCACTCCCAATAGCA 881
Qy      952 ATCAATTAGACCAATTTTGGTCTCTGCTACTATTATTAAGTCTTAAATT 1011
Db      882 CCGACATTTAGCAACATCTCACATTTCTTAACCTTGCCCTGGCCGATTTCTACTCT 941
Qy      1012 TCGGATATTA 1021
Db      942 TCTAATATA 951

```


QY 831 ATGCATTACTTACTTACTCAAGCTTGGTACCTTCACATGAACCTGGAAACGACC 890
 Db 73 AGCGGATGCTCGGATTTTCCACACTTGGATTCTTGAACCACTGGATCGACCTAC 14
 QY 891 TTGCAGA 897
 Db 13 GCGGCGA 7

RESULT 14
 AB020755 81672 bp DNA linear PLN 14-FEB-2004
 LOCUS Arabidopsis thaliana genomic DNA, chromosome 5, p1 clone:MZNI.
 DEFINITION AB020755 BA000015
 ACCESSION AB020755.1 GI:3985958
 VERSION
 KEYWORDS
 SOURCE Arabidopsis thaliana (thale cress)
 ORGANISM Arabidopsis thaliana
 REFERENCE 1 Seto, S., Nakamura, Y., Kaneko, T., Katoh, T., Asamizu, E., Kotani, H. and Tabata, S.
 AUTHORS Structural analysis of Arabidopsis thaliana chromosome 5. X.
 TITLE Sequence features of the regions of 3,076,755 bp covered by sixty
 JOURNAL DNA Res. 7 (1), 31-63 (2000)
 PUBMED 10718197
 REFERENCE 2 (bases 1 to 81672)
 AUTHORS Nakamura, Y.
 JOURNAL Direct Submission
 COMMENT Submitted (03-DEC-1998) Yasukazu Nakamura, Kazusa DNA Research
 Institute, Department of Plant Gene Research, 1532-3, Yana,
 Kisarazu, Chiba 292-0812, Japan (E-mail: ynakamu@kazusa.or.jp,
 Tel:81-438-52-3335, Fax:81-438-52-3334)
 Address for correspondence: kaos@kazusa.or.jp
 For the latest information on annotation of this clone, please see
 http://www.kazusa.or.jp/kaos/cgi-bin/agg_graph.cgi?c-MZNI
 Genes with similarity to proteins in the databases are described in
 'product' or 'note' qualifiers. Genes that have no significant
 protein similarity are described as 'unknown protein'.
 The software programs used to predict genes include: Graal
 (Informatics Group, Oak Ridge National Laboratory,
 http://compbio.ornl.gov/Graal-1.3/).
 GENSCAN (Chris Burge, MIT, http://CCR-081.mit.edu/GENSCAN.html),
 NetGene2 (S.M. Hebsgaard, et al., CBS, Technical University of
 Denmark, http://www.cbs.dtu.dk/services/NetGene2/) and
 SplicePredictor (Volker Brendel, Stanford University,
 http://gremjini.zool.iastate.edu/cgi-bin/sp.cgi).
 Genes encoding tRNAs are predicted by tRNAscan-SE
 (Sean Eddy, Washington University School of Medicine, St. Louis,
 http://genome.wustl.edu/eddy/tRNAscan-SE/).
 This sequence may not be the entire insert of this clone. It may be
 shorter because we remove overlaps between neighboring submissions.
 The 5' clone is MQ22 and the 3' clone is K19M22.
 location/Qualifiers
 1. 81672
 /organism="Arabidopsis thaliana"
 /mol_type="genomic DNA"
 /db_xref="taxon:3702"
 /chromosome="5"
 /clone="MZNI"
 /clone_lib="Mitout P1"
 /ecotype="Columbia"
 complement(join(49..299,375..684,765..1056,1148..2796))
 /note="unamed protein product; gb|AA14404.1
 gene_id:MZNI.1
 strong similarity to unknown protein"
 /codon_start=1
 /evidence=not experimental
 /protein_id="BAA97325.1"

CDS
 /db_xref="GI:8843777"
 /translation="MLYEOYAVLRSLGDAKLRLRCLNARRHAIDVPCRVESAMLR
 FFRBEHLVGVESMDGNGLASBCEPKTSLTGCDLNTVDDESCGCEGSDVDVI
 SKADEPGLDEVSDISFCVGSERAKKVSIRIALSPFEMALGSPFESTTSIDFB
 NGISIRAMALINIVSRILKRVDELFRVETVELILOAKFCDDILKREEARLASYDL
 DRALTVEYALERTTLILSACIQVLRLEIPSLNHPKVRPFCCSSAKQOLPIGSE
 CVELLYFLSIVQGMERKLTDTMLILERTREARNTWQALSLHQNGVLFERDK
 AAGFRLASLGHVSLAGVSRTEYKQGRVAYRLNMLISNHHKGMVYOBRSY
 NVGEKLDKDLATELDPKLSFPYKYRAVWPKFOKQI KEAFQEBIDRLIQPLSPECIE
 LRAMLVIATGDRSCGLDRALVLSLEPNVYFPGKRDLDLALTCQIEVESADCM
 VRLPDRSAVDVAVSLAVVHOMQNDPSKRYLFRQSLILRLNCGEAMRCILRMAN
 LATSCEARLYVSGMLIYDMGYVETITKAEALISIQSPSEAPFLKAYALDKULDE
 ISCVQVLEALKCPSDGLRKQGLNNLSITYNLQDLDAETAKALITKIRARQ
 GLARVFLNQRKACEEMTKLIEKSCSAAYEKSEYCEERAKEDLDMATTLDEL
 RTYPRYRAVIMDDQRETEAVELSKALAFRELEQTLHRAAFHEATGMVSLATDPC
 EALCDPNHTEHLIYSRSKDOASSIDNTIFELD"
 join(5111..5425,6229..6296,6383..6514,6654..6769,
 6876..7057,7543..7675,8596..8696,9202..9513,9628..9822)
 /note="unamed protein product; emb|CAB85585.1
 gene_id:MZNI.2
 similar to unknown protein"
 /codon_start=1
 /evidence=not experimental
 /protein_id="BAA97326.1"
 /db_xref="GI:8843778"
 /translation="MAITSTYTKLSVLCSSFISSPLVDSPSLAFSPIDRFLTVRIA
 TSFRSSSRPATIKRKSILAAVFPENSVISDVCAQVTSIVAFSCIFMGEGIKRGI
 FDOXLRKIVHINIGLVPMICPLFSSIGQALFASIPGLNIVRMLLGLGVHDYGR
 TIKMSRHGRDRELKGPLIYVLSITSAICYKRSPIAIVICNLACAGMDIVGR
 RRTETELPYNKKSPAGSIGMATAFGALVAVYVSPASRGVITRDSGMLIRPLVGLVT
 VIANFQTRKHQEBDEADDKYRDLTMMGVYSSEIMTIGFPFADVKFNRSEKRS
 LSLPPFVATRLVLSLSQIFLDVDSVLADVASECHRVARLGLDRDLVEELRLS
 VEARAKIADPSNNLENTKXVVDIFCGTHPPVASEVNCNRCORQIYAGFPAHLEK
 MGKGRARAKTTRSTTAQQRNARARNPNRPSYPYPSASBNQLASGSPGVAAGDCSNF
 TVRENVKGD"
 complement(11504..12430)
 /note="unamed protein product; gene_id:MZNI.3
 unknown protein"
 /codon_start=1
 /evidence=not experimental
 /protein_id="BAA97327.1"
 /db_xref="GI:8843779"
 /translation="MSBEDGSMGVSKSSSLFKIISVNSVTLAAVLVLLVLFV
 LILHFAFPFPMSSHQDFSAARHRRRRRRRTVYTRTIIISLPGLGGVDGSSPA
 ATKTRDCKGLSDSVISIPLFVEYENBEEDEBEVCITGLWEAGFGPKRLNCGG
 FYVEICIDMWLSHSTCPLCRSPVLAIVSDENKILANVAEEBAVLQSPAGENS
 NVSGDRVLSLSVMEDLKTGDDDEEYRIEVPDDDEIINGTSDRRRSWMTS
 SAASSLRMLSSSSSRKRVPTTARQDSK"
 complement(join(14616..14705,14920..15146,15667..15792,
 15880..16031,16161..16312))
 /note="unamed protein product; contains similarity to
 small GTPase Ran binding protein 1
 gene_id:MZNI.4"
 /codon_start=1
 /evidence=not experimental
 /protein_id="BAA97328.1"
 /db_xref="GI:8843780"
 /translation="MASTPERENRREDTEVNEDEDTGAQVAPIVRLAEVAVTTEED
 EDGAVLDLKSMTVPDKGNQMGKAGVATLGHKEKGYRLVNRQSKTLICANHLI
 SSGMSYQHSNGKSKCLMHTDPSDELDNGLCIRPAGSIENCTPMEKTEIAESQ
 VGRSTQGDENAGLIENTSVENISEKAEABEKPAKEDKTKKEVEBEKTEAS
 TIIVFGLKIVLSLSSSIIIRNGFVASRLAHN"
 complement(17150..17338)
 /note="unamed protein product; gene_id:MZNI.5
 unknown protein"
 /codon_start=1
 /evidence=not experimental
 /protein_id="BAA97329.1"
 /db_xref="GI:8843781"
 /translation="MTRKVPFDSINIQYRAANKSGSRVHEMMLYQSPFISINRCF
 GGNILLPIDCYRTSGSTL"
 complement(join(18542..18878,18974..19132,19206..19384,
 19467..19638,19916..20277))

CDS

THIS PAGE LEFT BLANK

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 16, 2006, 15:57:46 ; Search time 8511 Seconds
(without alignments)
5997.498 Million cell updates/sec

Title: US-10-634-548-1
Perfect score: 1091
Sequence: 1 aaaaagaagaataactacaa.....ccgacctacaagtcttc 1091

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: gb_eest1:*
2: gb_eest2:*
3: gb_eest3:*
4: gb_eest4:*
5: gb_eest5:*
6: gb_eest6:*
7: gb_eest7:*
8: gb_eest8:*
9: gb_eest9:*
10: gb_eest10:*
11: gb_eest11:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	631.4	57.9	636	1	AU236980 AU236980
2	507	46.5	507	6	CB261834 85-B864-
3	501	45.9	524	6	CB261126 26-B9406-
4	448.8	41.1	483	1	A1995595 701675020
5	415	38.0	418	1	AU227991 AU227991
6	314.4	28.8	650	7	CO981750 GM89012A2
7	313.8	28.8	864	7	DR914339 EST110587
8	307.8	28.2	881	7	CK286787 EST749509
9	278.8	25.3	787	6	CA297624 SCCSD2C0
10	265.6	24.3	654	2	CG412580 OV2.36.B1
11	264.2	24.2	734	8	DN207999 MEST864_B
12	256.4	23.5	883	7	CY711742 UCRPT01_0
13	255	23.4	675	8	DN229569 MEST1065
14	251.2	23.0	658	8	CD405855 GM-CK2993
15	237.8	21.8	690	6	CA277256 SCASD201
16	237.4	21.8	623	1	A1712188 60506508
17	233.6	21.4	650	4	A17109900 Zee may
18	232.4	21.3	788	10	CT508895 SAIL.804
19	227.2	20.8	680	8	DN212311 MEST7936_D
20	220.8	20.2	660	8	CK936854 CGF100451
21	211.2	19.4	690	1	AU056832 AU056832
22	209.8	19.2	508	7	CK748049 ltu01-18m

C	23	207.8	19.0	596	6	CD392779	GM-CK1203
C	24	202.6	18.6	898	8	CK113647	ET083D19
C	25	197	18.1	941	7	CK297713	EST760427
C	26	196	18.0	569	1	A1833732	605094E10
C	27	191.4	17.5	640	1	AW735812	EST336580
C	28	190.6	17.5	623	2	BG412891	OV2.36.B1
C	29	189.6	17.4	570	6	CA820273	CA866905
C	30	183.4	16.8	444	3	BM143737	sa148D08
C	31	183.2	16.8	546	6	CA202669	SCSGFL108
C	32	179.2	16.4	897	8	DR914340	EST110587
C	33	175.6	16.1	634	3	BM080425	MEST108-C
C	34	169	15.5	489	2	BG840419	MEST12-B0
C	35	169	15.5	489	2	BG840843	MEST12-B0
C	36	169	15.5	743	8	DN208011	MEST864_D
C	37	168.4	15.4	735	8	CM206873	Tor7293-G
C	38	168.2	15.4	422	9	BH701821	BOMMD17F
C	39	168.2	15.4	824	9	BH556465	BOGWI9TR
C	40	166.2	15.2	451	6	CD410254	GM-CK3786
C	41	166.2	15.2	462	5	BQ536471	STEM4.1.D
C	42	166	15.2	483	2	BE125719	DG1.54.B0
C	43	165.8	15.2	830	8	CY760662	FGA505E04
C	44	164.4	15.1	1125	8	DR740330	FGA508E25
C	45	162	14.8	480	6	CA820469	sa89f01

ALIGNMENTS

RESULT 1
LOCUS AU236980 636 bp mRNA linear EST 01-APR-2002
DEFINITION AU236980 RAPL15 Arabidopsis thaliana cDNA clone RAPL15-38-N13 5',
ACCESSION AU236980
VERSION AU236980.1 GI:19876149
KEYWORDS mRNA sequence.
SOURCE Arabidopsis thaliana
ORGANISM Arabidopsis thaliana (thale cress)
EST.
Arabidopsis thaliana

Arabidopsis thaliana
Bakaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE

1 (bases 1 to 636)

Seki, M., Narusaka, M., Ishida, J., Kamiya, A., Saito, M., Nakajima, M.,
Akiyama, K., Enju, A., Oono, Y., Sakurai, T., Carninci, P., Kawai, J.,
Itoh, M., Ishii, Y., Arakawa, T., Shibata, K., Shinozaki, K.,
Miyamatsu, M., Hayashizaki, Y. and Shinozaki, K.
Large scale analysis of Arabidopsis full-length cDNA
Unpublished (2002)

COMMENT

Contact: Motoaki Seki
Plant Functional Genomics Research Group
RIKEN Genomic Sciences Center
3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
Tel: 81-298-36-4359
Fax: 81-298-36-9060
Email: msek@rc.riken.go.jp
An Arabidopsis full-length cDNA library was constructed essentially
as reported previously (Seki et al., 1998). cDNA cleaved with BamHI
and XhoI was ligated to modified Lambda FIC-1 vector (Carninci et
al., submitted for publication) digested with BamHI and SalI. This
clone is in a modified pBluescript vector. Please visit our web
site (http://www.gen.riken.go.jp/e/plant/index_e.html) for further
details.

FEATURES

source

location/Qualifiers
1..636
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/clone="RAPL15-38-N13"
/tissue_type="mixture of silique and flower"
/lab_note="DH10B"
/clone_lib="RAPL15"
/note="Site_1: BamHI; Site_2: SalI"

ORIGIN

FEATURES
source

Location/Qualifiers
1. .864
/organism="Aquilegia formosa x Aquilegia pubescens"
/mol_type="mRNA"
/db_xref="taxon:338618"
/clone="COLID02"
/tissue_type="mixed shoot and floral apical meristems,
flower buds, leaves and roots"
/lab_host="DH10B T1 (T1 and T5 phage resistance)"
/clone_lib="Aquilegia cDNA library"
/note="Vector: pCMV SPORT6.1; Site 1: EcoRI, Site 2: NotI;
F2, F3, and F4 lines of Aquilegia Formosa x A. pubescens
were grown from seed in greenhouses at UC Santa Barbara.
From these plants three sets of tissue were collected: 1)
Small flower buds (<10 mm) and very young inflorescences
(71 & 29% by weight respectively), 2) Medium (7-20 mm) and
large (at or near anthesis) flower buds (65 & 35% by
weight respectively) and 3) Shoot apical meristems. A
fourth set of tissue was collected from plants of A.
formosa. These plants were grown from seed in sand and at
approximately 1 month root tissue and leaf tissue of
various developmental stages were collected (84 & 16% by
weight respectively). Total RNA was extracted from each
set of tissue and pooled in the following proportions:
1.5X from sets 1 & 2, 1X from sets 3 & 4. From the pooled
total RNA, mRNA was extracted and enriched for full-length
messages and then normalized with proprietary methods by
Invitrogen."

ORIGIN

Query Match 28.8%; Score 313.8; DB 8; Length 864;
Best Local Similarity 64.9%; Pred. No. 1e-76;
Matches 465; Conservative 0; Mismatches 252; Indels 0; Gaps 0;

304 CGCTGTCAGAGAGTCGAGACCCGCGGAGAGCTTGGAGAGCATCCCGCTTGTCT 363
Db CACTTGGCAGAGTGCAGAGTCGGCGGACCTGTAGTGGTGGCTACGCTCTTGTAC 732
364 TAAAGTTCGAGAGTCTCAAGCGAAAGTCATTCACAGAGTTTGAGCAAGAGCTTG 423
Db GTTCTTTGATATCTTACCTAGAGGATTTATCAACAGAGTTTGAGCAAGAGCTTG 672
424 TGCATATCTCTCAGAGTCTGCTTTTCTGACTTCGCTGCGCAATCTTCAGCGAATCGACG 483
Db TTCAATATGTCTGGGCTCTATTCATGCTTCTTGCGCAATTTTATGTAATCAACCG 612
671 TTCAATATGTCTGGGCTCTATTCATGCTTCTTGCGCAATTTTATGTAATCAACCG 612
484 AGGCTCCATCTTGGCTTTTGTCCGTTAGAGAGGCTTAAAGCTTGTATTAAG 543
Db AAGCTCCATCTTGGCTTTTGTCCGTTAGAGAGGCTTAAAGCTTGTATTAAG 543
611 AAGCTCCATCTTGGCTTTTGTCCGTTAGAGAGGCTTAAAGCTTGTATTAAG 552
544 GACTATCCATCTTCCCAATTCAGATGTAATCAATCCGTCACAAGAGAGAGAGAG 603
Db GACTATCCATCTTCCCAATTCAGATGTAATCAATCCGTCACAAGAGAGAGAGAG 603
551 GACTTTCATGAGAGAGTCTCTCTACTACTGTTTGAATCTGTACACAGTGAAGAGATCCA 492
491 AGGAATCTCTGAGAGAGTCTCTCTACTACTGTTTGAATCTGTACACAGTGAAGAGATCCA 492
604 AAGAGTTCCTTAAAGTCTTTGTTTCTAGCTTCTTCTTTCTTCTGCGGTTTTCT 663
Db AAGAGTTCCTTAAAGTCTTTGTTTCTAGCTTCTTCTTTCTTCTGCGGTTTTCT 663
491 AGGAATCTCTGAGAGAGTCTCTCTACTACTGTTTGAATCTGTACACAGTGAAGAGATCCA 432
664 TCTGAGAGAGTCTCTCTGATGATCTGTTAGCAATGATGTGTGTGGCCATGAGAA 723
Db TCTGAGAGAGTCTCTCTGATGATCTGTTAGCAATGATGTGTGTGGCCATGAGAA 723
431 TCTGAGAGAGTCTCTCTGATGATCTGTTAGCAATGATGTGTGTGGCCATGAGAA 372
724 TAGCTGATATATAGGAGATGAGTGGTCACTTAAGATACCTTACCAACCAAGAAAGA 783
Db TTGCTGATATATAGGAGATGAGTGGTCACTTAAGATACCTTACCAACCAAGAAAGA 312
371 TTGCTGATATATAGGAGATGAGTGGTCACTTAAGATACCTTACCAACCAAGAAAGA 312
784 GTTGGGAGAGAGATCTCATGTTTCTTCTGCTTCTTCTCATCTCATGCAATTAATTT 843
Db GTTGGGAGAGAGATCTCATGTTTCTTCTGCTTCTTCTCATCTCATGCAATTAATTT 843
311 GTTAGCTGAGAGATTTCTCATGTTTCTTATTTGGTTCTTGTCTCATTCATGTTGCTAT 252
844 ACTATTTACTCAAGCTTGGGATCTTCAATGAACTGGGAAAGCACTTGGAGAGAGTAG 903
Db ACTATTTACTCAAGCTTGGGATCTTCAATGAACTGGGATCTTCAAGCTGGGAAAGTAG 192

904 CAATGTCATAGTGTGCGACGATGCTGAGTCCAGTACCCATCAACCATTAAGACG 963
Db CTTTAGTAGCTTGTGTGGCAACATAGTAGAGTCCCTTCTTACTACAGAGATAGTAGATG 132
964 ACAATATTTGGTGTCTCTGCGTCACTATTTTACTGCTTATTTAAGTTTGGATATT 1020
Db ACAATATTTGGTGTCTCTGCGTCACTATTTTACTGCTTATTTAAGTTTGGATATT 75

RESULT 8

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

OTHER ESTS

CONTACT

INSTITUTE

SEQ PRIMER

81 bp mRNA linear EST 02-AUG-2004
EST749509 Nicotiana benthamiana mixed tissue cDNA library,
normalized, full-length Nicotiana benthamiana cDNA clone NBM462.5'
end, mRNA sequence.
CK286787 GI:39862689
EST.
Nicotiana benthamiana
Nicotiana benthamiana
Bakryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
asterids; lamiales; Solanales; Solanaceae; Nicotiana.
1 (bases 1 to 881)
Buell, C.R., Hart, A., Zismann, V., Karamycheva, S.A., Day, B.,
Staskiewicz, B., Jin, H. and Baker, B.
Generation of EST sequences from Nicotiana benthamiana
Unpublished (2003)
Other ESTs: EST749510
Contact: Robin Buell
The Institute for Genomic Research
9712 Medical Center Dr, Rockville, MD 20850, USA
Email: potlato-array@tigr.org
Clones can be requested from the University of Arizona Genomics
Institute via <http://genome.arizona.edu/orders/>.
Seq primer: ATT TAG GTG ACA CTA TAG.

FEATURES

source

Location/Qualifiers
1. .881
/organism="Nicotiana benthamiana"
/mol_type="mRNA"
/db_xref="taxon:4100"
/clone="NBM462"
/tissue_type="abiotic and biotic stress-treated leaves,
callus tissue and root tissue"
/lab_host="DH10B-TonA"
/clone_lib="Nicotiana benthamiana mixed tissue cDNA
library, normalized, full-length"
/note="Vector: pCMVSPORT6.1; Site 1: EcoRI, Site 2: NotI;
supplier: RNA was isolated from Nicotiana benthamiana
tissues that include callus, roots from liquid culture
grown plants, heat-stressed leaves (38 C, 3 hr and 6 hr),
cold-stressed leaves (5 C 3 hr, 6hr), and pathogen
challenged leaves (Pseudomonas syringae pv tomato 12 hr;
Xanthomonas campestris pv campestris 12 hr, 18hr;
Pseudomonas syringae pv phaseolicola 18hr, and Xanthomonas
campestris pv vesicatoria 18hr). RNA was isolated from
these tissues and pooled in approximately equal molar
amounts."

ORIGIN

Query Match 28.2%; Score 307.8; DB 7; Length 881;
Best Local Similarity 67.2%; Pred. No. 5e-75;
Matches 435; Conservative 0; Mismatches 212; Indels 0; Gaps 0;

303 TCGCTGTTCATGACGTCGAGACCAAGTGGTGTGAGAGATAGCGCTTGTCTC 362
Db TCGCTGTTCATGAGATGCGGAGACCAAGTGGTGTGAGAGATAGCGCTTGTCTC 294
363 TTAAGCTTCGAGAGTCTCACCAAGCGAAGTCAATTAAGAGATTTGAGCAAGAGCTT 422
Db TCAAGCTTCGAGATTTCTTATTCGAGCGAAGCTCATTAAGAGATTTAAGCAAGAGCTT 354

443 GGCACATTAATCTCAAGAGTGTGGCTTTTCGTAATTGCGTGGCCAAATCTTCAGCGGATCGAAC 482

Db 355 GTCCACATTAATGTCTGGGTCTGGCTTTTATATAGCGTCTCTGGCCAAATTTTCAGTCAATCAGAA 414

Qy 483 GAGGCTCGATACTTTGCTGCTTTTGTTCGGTATAGTAATGAGCTTAAGGCTTTGTATTAAC 542

Db 415 TGGGACACGTACTTTTGCTTCTGTAGTTCGGCTTAACAAATGTATTGACTTTGATTTAT 474

Qy 543 GGACTATTCATTTCCCCCAATTCGATGCTAATCAAAATCCGTCAACAAGAGAGGAGAGA 602

Db 475 GGCCTTTCTTTGGCTACGATAGAGGACCTTCMAAATCTGTACTACGGAGAGAAAGCA 534

Qy 603 GAAGAGTGTCTAAAGGTCCTTTGTTCACGTTCAAGCTCTCTTTTCTCTGGGTTTTTC 662

Db 535 GAAAGAAATGTCTTAAGAGGCTCTTAATTAATATGTTCTTAGTGTAAATTTGACCCGACTTCTC 594

Qy 663 TTCTGAGAGAGTCTCTTAACGGTATGATCTCGTTAGCAATGATGTGTGGTGCATGGA 722

Db 595 TTTTGGCGTAGTCAACGATGGAGTAATTTGTTAGCAATATGTGTGTGTATGGA 654

Qy 723 ATAGCTGATATATAGGACGTAAAGTTTGGGTCACTAAGATACCTTACCAACCAAGAAAG 782

Db 655 ATTGTCTGATATGTTGAAAGAGTTTGGGTCCATTAATAATCCCTTAATATAACAGAAA 714

Qy 783 AGTTGGGCAAGAACATCTCCATGTTCAATCTTGGCTTCTTATCTTCATGCACTTATCTT 842

Db 715 AGTTTGGGTGGTAGGCTCTCCACATGTTTGTTCCTCGTTTCTGTGTGTCAATTTGGAGATGCTC 774

Qy 843 TACTATTAATCAAGCCTTGGGTACCTTCACATGAATGGGAAACAGACCTTGAGAGAGTA 902

Db 775 TATTACTTCTCTCTCTTGGGAATATCTTCACTTGAATGGGTTTCAACGTATAAAAGATGA 834

Qy 903 GCAATGTCTCAATGATGCGCCACGGTAGTCAATGCGTCAACCATCAC 949

Db 835 GCTGTGTGTCAATTATAGCACTATGATGGAGTCTTAACCTAATTAC 881

FEATURES	source
RESULT 9	
LOCUS	CA297624/c
DEFINITION	CA297624, 787 bp, mRNA, linear, EST 26-SRP-2002
ACCESSION	SCCCSDC02H02.g SD2 Saccharum officinarum cDNA clone SCCCSDC02H02
VERSION	5', mRNA sequence.
KEYWORDS	CA297624
SOURCE	CA297624.1 GI:36066481
ORGANISM	EST.
	Saccharum officinarum
	Saccharum officinarum
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
	SpERMATOPHYTES; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCD
	clade; Panicoideae; Andropogoneae; Saccharum; Saccharum officinarum
	complex.
REFERENCE	1 (bases 1 to 787)
AUTHORS	Vecitore,A.L., da Silva,P.R., Kemper,E.L. and Arruda,P.
TITLE	The 1libraryes that made SUCEST
JOURNAL	Genet. Mol. Biol. 24 (1-4), 1-7 (2001)
COMMENT	Contact: Arruda P
	Centro de Biologia Molecular e Engenharia Genetica
	Universidade Estadual de Campinas
	Caixa Postal 6010, 13083-970, Campinas SP, Brazil
	Tel.: 55 19 3788 1137
	Fax: 55 19 3788 1089
	Email: parrruda@unicamp.br
	Clone distribution: clone distribution information can be found
	through the Brazilian Clone Collection Center (BCCC) at
	http://www.bccccenter.fcav.unesp.br
	Plate: C02 row: H column: 02
	Seq primer: T7 Promoter Primer.
	Location/Qualifiers
	1..787

FEATURES	SOURCE	location/qualifiers
		1..787
		/organism="Saccharum officinarum"
		/mol_type="mRNA"
		/db_xref="taxon:4547"
		/clone="SCCSD2C02H02"

`/lab host="DH1.0B"
/clone lib="SD2"
/note="Organ: Developing seeds (small insert library);
Vector: pSport1; Site 1: SalI; Site 2: NotI; An
unidirectional cDNA library generated from [Developing
seeds (small insert library)]. cDNA was prepared from
poly(A+ mRNA using Superscript Plasmid System Kit
(Invitrogen). The double-strand cDNAs were fractionated
in a sepharose CL-2B 40cm-columns and fragments sizing
between 0.8 and 1.5 kb were directionally cloned into the
vector. Details of each source of RNA and library
construction can be obtained at
http://sucrest.lad.ic.unincamp.br/pub/ldc"`

[illegible]

RESULT 10	LOCUS	DEFINITION
BG412580	654 bp	mRNA
BG412580	linear	EST 13-MAR-2001
OV2_35-B1.1.g1_A002	Ovary 2 (OV2)	Sorghum bicolor CDNA, mRNA sequence.

ACCESSION BG412580
 VERSION BG412580.1 GI:13318133
 KEYWORDS EST.
 SOURCE Sorghum bicolor (sorghum)
 ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Sorghum.
 REFERENCE 1 (bases 1 to 654)
 Cordonier-Pratt, M.-W., Gingle, A., Marsala, C., Sudman, M. and Pratt, L.H.
 TITLE An EST database from Sorghum: ovaries of varying immature stages
 JOURNAL Unpublished (2000)
 COMMENT Laboratory for Genomics and Bioinformatics
 The University of Georgia, Department of Plant Biology
 Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
 Tel: 706 542 1860
 Fax: 706 583 0210
 Email: mmpratt@uga.edu
 Seg primer: PolyTmix
 High quality sequence start: 8
 High quality sequence stop: 628
 POLYA=No.
 FEATURES
 source
 1..654
 /organism="Sorghum bicolor"
 /mol_type="mRNA"
 /clone_lib="Ovary 2 (OV2)"
 /note="Organ: Mix of ovaries of varying immature stages from 8-week-old plants; Vector: pluscript II from lambda zap II; Site 1: XhoI; Site 2: EcoRI; The library was made from poly-A RNA in the cloning vector lambda zap II. Clones to be sequenced were prepared by mass excision."
 ORIGIN
 Query Match 24.3%; Score 265.6; DB 2; Length 654;
 Best Local Similarity 64.1%; Pred. No. 3.5e-63;
 Matches 400; Conservative 0; Mismatches 224; Indels 0; Gaps 0;
 Db
 368 CTTCGAGAGTTCACCAAGCAAGCGTATTCACAGAGTTGAGCAGAAAGCTTGCA 427
 3 CTTCGACAGAGTTCACCGAGCGGCGCTCATCGAAAGGTTGAGCAGAAAGTTGCA 62
 428 TATCTCTCAGTCTGCTTTCGTACTTCGCGGCATCTTCAGCGGATGACCGAGGC 487
 63 CGTGTATTCGAGCGCTCTGTTCACTGCTTGGCGCTGTTTCAAGCAATTCAGCAGAGC 122
 488 TCGATACTTTCGCTTTCGTTCGTTAGTGAATGGCTTAAAGCTTGTATTAACGACT 547
 123 ACGGATTCGCTGAGTTGTTCCACTCTTGAACCTCAATTAAGGCTTGTATTAAGACT 182
 548 ATTCATTTCCCAATTTGATGCTATCAATTCCTCCAGAGAGAGAGAGAGAGAGA 607
 183 CCGTCTCTACACTGATGAAGCTTAGTAAATCACTGACAGTGAAGAGAAACAGAGGA 242
 608 GTTGTAAAGGCTTTCGTCTAGCTTCGCTTCTTTCCTGCGGTTTCTTCTTG 667
 243 ATTGTGAGAGGCTCACTGATTAATGTTGCTGCTGTTTCACTGTTTATTAAGTCTTG 302
 668 GAGAGAGTCTCTATCGGTATGATCTCGTTAGCAATGATGTGTGCGATGAGATAGC 727
 303 GCGGAGAGTCCCTGTTGGGATGTTCTTTTGTCAATGATGATGTGTGCGATGAGTTCG 362
 728 TGATTAATGGAGCTAAGTTGGGTCACTAATGATACCTTAACAACCAAGAGAGTTG 787
 363 TGACATTTGTGGAGAGGATGAGCTCAGTGAAGCTGCATTCATTAAGAGAGAGCTG 422
 788 GGCAGAGAGCATCTCATGTTGATCTTGCGGCTTCTTCAATCCATCGCATTAATTACTTA 847
 423 GCGCGGAGCATCTTAATGTTCAATTTCTGCTTCTGCTGCGGATGATGATGTTCTTA 482

Qy 848 TTAATCAAGCTTGGGATCCTTCATGAACTGGGAGAAACGACTTGACAGAGTACCAAT 907
 Db 483 CTTCGAGAGCTTGGTACATGATGTTATTTGGCAGAGGCACTTGTAAGCTGCGCT 542
 Qy 908 GGTCTCAATGTCGCCACGGTACTGAGTGCATACCATCAAGCAATTAAGCAGCA 967
 Db 543 TGTTCACCTGGCAGCAACCGTATGAGTATTCCTGATCAGAAAGTTAGATGACAA 602
 Qy 968 TATTCGCTTCCTGCTGCTACTAT 991
 Db 603 TATATCTGTTCTTGGCACCACAT 626

RESULT 11
 LOCUS DN207999/c 734 bp mRNA linear EST 28-FEB-2005
 DEFINITION MEST864.B02.T7-1 UGA-zmsam-XZ2 Zea mays cDNA, mRNA sequence.
 ACCESSION DN207999
 VERSION DN207999.1 GI:60341026
 KEYWORDS EST.
 SOURCE Zea mays
 ORGANISM Zea mays
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.
 1 (bases 1 to 734)
 Chen, H.D., Zhang, X., Zhou, R.L., Arias, L.A.C., Shendelman, J.M., Zazubovits, N., Boreuk, L.A., Emrich, S.J., Ashlock, D.A., Scanlon, M.J. and Schnable, P.S.
 Expressed Sequence Tags from B73 Maize Shoot Apical Meristems
 Unpublished (2004)
 CONTACT: Patrick S. Schnable
 Schnable Laboratory
 Iowa State University
 2035B R.J. Carver Co-Lab, Ames, IA 50011-3650, USA
 Tel: 515-294-0975
 Fax: 515-294-5256
 Email: schnable@iastate.edu.
 location/Qualifiers
 1..734
 /organism="Zea mays"
 /mol_type="mRNA"
 /cultivar="Indred B73"
 /db_xref="taxon:4577"
 /tissue_type="Vegetative Shoot Apical Meristem (SAM) and leaf primordia staged P1-P4"
 /lab_host="XU1-Blue"
 /clone_lib="UGA-zmsam-XZ2"

FEATURES

source

/note="Organ: Shoot apex; Vector: Uni-Zap XR; Site 1: EcoRI; Site 2: XhoI; This library was constructed by Xiaolan Zhang. Vegetative Shoot Apical Meristem (SAM) and leaf primordia staged P1-P4 from 14-17 day-after germination seedlings were quickly dissected into dry ice under a light microscope. Total RNA was isolated using Trizol and mRNA was purified with Dynal Oligo-dT25. ds-cDNA molecules were generated as follows. First-strand cDNA was prepared from oligo-dT selected mRNA by priming with an XhoI oligo-dT primer
 (5'-GAGAGAGAGAGAGAGAGAGAGAGTCTGAGTTTCTTTTCTTTT).
 The resulting DNA:RNA hybrid was treated with RNase H and used as a template for DNA Pol-I-catalyzed second strand synthesis. After the addition of EcoRI adaptors, the ds-cDNAs were digested with XhoI and size-selected to be >600 bp. The resulting molecules were directionally cloned into the EcoRI and XhoI sites of the Uni-Zap XR vector. The lambda library was packaged with Gigapack III Gold packaging extract and was mass excised by XU1-blue cells and Exsist helper phage. Excised phagemids were titered in SOLR cells and plated onto LB-ampicillin agar plates. Base calling was conducted using Phred. Trimming was performed using Lucy with the following criteria:
 (-minimum 200-error 0.01 0.01-bracket 10 0.01). A low complexity filter was applied and additional trimming was

conducted to remove B. coli, vector, and organelle contamination. After processing ~10% of the sequences contained a minimum of 10 Ts at the beginning of the sequence. For reasons that are not understood many of the clones in this library lack an XhoI site at their 3' ends."

ORIGIN

Query Match 24.2%; Score 264.2; DB 8; Length 734;
 Best Local Similarity 63.8%; Pred. No. 9e-63;
 Matches 401; Conservative 0; Mismatches 228; Indels 0; Gaps 0;

```

Qy 378 CTCACCAAGCGAAAGCTATCAACAGAGTTGAGCAGAAAGCTGTCATATCTCA 437
Db 728 CTCACCAAGCGCGCGCTATGAAAGAGCTTGGACAGAAAGTTGTGACGTGCTCC 669
Qy 438 GGTCTGCTTTTCGATCTTGCGGCAATCTTACGCGAGTCGACCGAGCTCATCTT 497
Db 668 GCGCTCTGTTCAATGTCATCTTGCGGCTGTTGAGCAATTCAGAGAGCAAGGATTTTC 609
Qy 498 GCTGCTTTTTCGTTAGTGAATGCTTAAAGGCTTGTATTAACGACTATCCATTTC 557
Db 608 GCGCGGTTGTCGCTTCTGAACTCATGAGGCTTCTGATATATGAGCTCCGCTCTAC 549
Qy 558 CCAATTCGATGCTATCAATCCGTCACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 617
Db 548 ACTGATGAGCTCTGTTAAATCACTGACAGTGAAGAGAAACAGAGAAATGCTGAGA 489
Qy 618 GGTCTTTTTCGATCTGAGCTTCTTCTTCTGCGGTTTCTTCTGAGAGAGTCT 677
Db 488 GGTCTCTATATATGCTGCTGCTGCTGTTGAGCGGTTTATGCTTCTGCGAGAGTCC 429
Qy 678 CTTATCGGTATGATCTGTTAGCAATATGTTGTGTCGCGAGTGAATGCTGATATATG 737
Db 428 CCGATCGGAGTCTCTCTCTGATGATGAGCGGTGCGAGTGTGCTGACATGTT 369
Qy 738 GAGAGTATGTTGGGTCATTAAGTATCTTACCAACCAAGAGAGTGGGAGAGAGC 797
Db 368 GGGAGAGTATGCTGACGAGAGCTGCAATTCAGAGAGAGAGAGAGAGAGAGAGC 309
Qy 798 ATCTCATGTTTCATCTTGCGCTTCTTCAATCTGCAATTCATTAATTAATCAAGC 857
Db 308 ATCTCATGTTTCATCTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCAAGC 249
Qy 858 CTTGGGATCTTCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATG 917
Db 248 CTTGGGATCTTCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATG 189
Qy 918 GTGCGCAGAGTATGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 977
Db 188 GCGAGCAGAGTATGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 129
Qy 978 CTTCTGCTACTATTTAGTCTTATTT 1006
Db 128 CTTTGGCAGCAGAGTGTGCTTCTTCT 100

```

RESULT 12
 CV711742/c 883 bp mRNA linear EST 03-NOV-2004
 LOCUS UCBP701.0016005.f Poncirus trifoliata CTV-challenged cDNA library -
 DEFINITION AG12 Poncirus trifoliata cDNA clone PT_68A0016005, mRNA sequence.
 ACCESSION CV711742 GI:55294110
 VERSION CV711742.1
 KEYWORDS EST.
 ORGANISM Poncirus trifoliata
 SOURCE Poncirus trifoliata
 Bacteria: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophytes; Magnoliophyta; eudicotyledons; core eudicotyledons;
 rosids; eurosids II; Sapindales; Rutaceae; Poncirus.
 1 (bases 1 to 883)
 Rose, M.L., Ye, X., Federici, C.F., Close, T.J., Fenton, R.D.,
 Manamaker, S., Kim, H.R., Kudrna, D. and Stum, D., MIsosetski, M.,

TITLE
 JOURNAL
 COMMENT
 Ming, R.
 Development of EST Resources and New Genetic Markers for California
 Citrus - Poncirus trifoliata CTV-challenged phloem - AG12
 Unpublished (2004)
 Contact: Mikeal Roose
 Department of Botany & Plant Sciences, University of California
 Riverside, CA, 92521-0124, USA
 Tel: 9097874437
 Fax: 9097874437
 Email: mikeal.roose@ucr.edu
 Seq primer: T7.

FEATURES

source

Location/Qualifiers

1. 883

/organism="Poncirus trifoliata"

/mol_type="mRNA"

/cultivar="Pomeroxy Op"

/db_xref="taxon:37690"

/clone="PT_68A0016005"

/tissue_type="Phloem"

/dev_stage="10 - 30 cm shoots"

/lab_host="B. coli TUC121"

/clone_id="Poncirus trifoliata CTV-challenged cDNA library - AG12"

/note="Vector: Lambda Uni-ZAP XR, excised phagemid;
 Site 1: EcoRI; Site 2: XhoI; Plants were grown in the
 greenhouse at University of California, Riverside. The
 section was an open-pollinated (very probably selfed)
 seedling of Poncirus trifoliata cv Pomeroxy that was
 selected as homozygous for the CTV resistance gene. The
 rootstock was sweet orange infected with citrus tristeza
 virus (CTV) isolate T514 over 1 year before sampling (CTV
 infects sweet orange, but not genotypes carrying the CTV
 resistance gene). Shoots 10-30 cm long were harvested in
 October 2000, and the green phloem (dark) was removed and
 frozen quickly in dry ice. Total RNA was extracted using
 TRIzol reagent (Gibco). Poly(A) RNA was purified, a cDNA
 library was made, and 0.5 million primary lambda cDNA
 clones were in vivo excised to give a population of
 phagescript SK(-) phagemids. All steps to this point were
 performed in the ML Roose lab at the University of
 California, Riverside by X. Ye. Phagemids were plated,
 plasmid DNA purified, cDNA clones archived, and DNA
 sequences determined bi-directionally using an ABI3730 at
 the Arizona Genomics Institute, University of Arizona
 (Kim, Kudrna, Stum, MIsosetski, Ming). Chromatogram files
 were downloaded to UC Riverside (Close), then processed at
 UC Riverside (Manamaker) using the Harvest pipeline
 (http://harvest.ucr.edu) to remove vector and cloning
 oligo sequences and various contaminants, and to trim to a
 high quality region. Sequences that retained a phred 17
 region of at least 100 bases were deposited to Genbank."

ORIGIN

Query Match 23.5%; Score 256.4; DB 7; Length 883;
 Best Local Similarity 66.0%; Pred. No. 1.5e-60;
 Matches 371; Conservative 0; Mismatches 191; Indels 0; Gaps 0;

```

Qy 471 AGCGATGACCGAGGCTGATCTTGTCTTCTTCTGTTAGTGAATGCTTAAG 530
Db 703 AGCAGCTCAACAGAGGCTGCTACTTGTCCGCTTGTCTCTGTGAATGCTTGA 644
Qy 531 CTTGTATTTAAGGAGCTATTCATTTCCCAATTCATGCTATCAATTCGTCACAGA 590
Db 643 CTTGTATTTAAGGAGCTATTCATTTCCCAATTCATGCTATCAATTCGTCACAGA 584
Qy 591 GAAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 650
Db 583 GAAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 524
Qy 651 TCTGCGGTTTCTTCTGAGAGAGTCTCTATGCGTATGATCTGTTAGCAATGATGTG 710
Db 523 TCTGCTCTGTTCTTCTGAGAGAGTCTCTATGCGTATGATCTGTTAGCAATGATGTG 464

```

QY	711	GGTGGCAATGGAATAGCTGATATATAATGGAGCGTAAAGTTGGGTCAACTGAATACCTTAC	770
Db	463	GGGGAGATGTGTCTCTACCGTTRTCGGAGAAAGATTTGGGTCCATBAAGTCTTTTAT	404
QY	771	AACCCAGAAAAGATTGGGACAGAAAGCATCTCCATGTCATCTTCGGCTTCTTACATCC	830
Db	403	AATGAGAAAAAAGTTGGGTGGGACGATATCAATGTTTGTTTTGGATTCCTTAGTTTCC	344
QY	831	ATCGCAATTACTTACTTATTAATTAACAGCCTTGGGTAACCTTCACTAGAACTCGGGAAACGAC	890
Db	343	ATTGGAGATGCTGTACTTCTATTCGATTTTGGGATTAATTAACCACTGGAATTTATTTAGAC	284
QY	891	TTTGCAGAGATGACAAATGTCATAGTGTCCGACAGGATGTCGATGCTGTACCATCCATCACC	950
Db	283	TTTGCACCGGTGCTTTAGTGTCTCTAGATGGGAGAGTTGTGCAATCTCTCCCGATTACA	224
QY	951	GATCAATTAGACGACAAATATTTCGGTGTCTGTGGCTACTATTTTATAGCTGTATTAGT	1010
Db	223	GAGGTTGTGATGACAAATATATCTGTCTTCTCTTGCAGACATGTAAGCAGCAATTTTAGT	164
QY	1011	TTTCGATATTAGATTAAATCCCT	1032
Db	163	TTTGGTTATTAGATTGTACCTT	142

RESULT 13	LOCUS	DEFINITION
DN229569/c	DN229569	675 bp mRNA linear EST 01-MAR-2005
	MEST1065_F09.T7-1	UGA-2msm-Xx2 Zea mays cDNA, mRNA sequence.

KEYWORDS	EST.
SOURCE	Zea mays
ORGANISM	Zea mays

REFERENCE
AUTHORS
1 (bases 1 to 675)
Chen, H. D., Zhang, X., Zhou, R. L., Arias L. A. C., Shendelman, J. M., ...

TITLE	Expressed Sequence Tags from B73 Maize Shoot Apical Meristems
JOURNAL	Unpublished (2004)
COMMENT	Contact: Patrick S. Schnable

2035B R.J. Carver Co-Lab, Ames, IA 50011-3650, USA
Tel.: 515-294-0975
Fax: 515-294-5256
Email: ecmnhale@astate.edu.

FEATURES

source

```

/organism="Zea mays"
/mol_type="mRNA"
/cultivar="Inbred B73"
/db_xref="taxon:4573"
/feature_type="Vegetative Shoot Apical Meristem (SAM) and
leaf primordia staged P1-P4"
/lab_host="X11-Blue"
/clone_id="UGA-ZMSAM-X22"
/note="Organ: Shoot apex; Vector: Uni-Zap XR; Site: 1:
XbaI; Site 2: XhoI; This library was constructed by
KaoIan Zeng. Vegetative Shoot Apical Meristem (SAM) and
leaf primordia staged P1-P4 from 14-17 day-after
germination seedlings were quickly dissected into dry ice
under a light microscope. Total RNA was isolated using
Trizol and mRNA was purified with Dynal Oligo-DT25.
5s-cDNA molecules were generated as follows. First-strand
cDNA was prepared from oligo-dT selected mRNA by priming
with an XhoI oligo-dT primer
(5'-GACGAGAGAGAGAGAGAGAACACAGCTCTGAGCTTTT
TTTTTTTTTTTTTTT).
The resulting DNA:RNA hybrid was treated with RNase H and
used as a template for DNA PolI-catalyzed second strand

```

synthesis. After the addition of EcoRI adaptors, the ds-cDNAs were digested with XhoI and size-selected to be >600 bp. The resulting molecules were directionally cloned into the EcoRI and XhoI sites of the Uni-Zap XR vector. The lambda library was packaged with Gigapack III Gold packaging extract and was mass excised by λ LI-Blue cells and ExbaStar helper phage. Excised phagemids were titered in SOLR cells and plated onto LB-ampicillin agar plates. Base calling was conducted using Phred. Trimming was performed using Lucy with the following criteria: (-minimum 200 -error 0.01 0.01 -bracket 10 0.01). A low complexity filter was applied and additional trimming was conducted to remove E. coli, vector, and organelle contamination. After processing ~30% of the sequences contained a minimum of 10 Ts at the beginning of the sequence. For reasons that are not understood many of the clones in this library lack an XhoI site at their 3' ends."

ORIGIN

Query Match	23.4%	Score 255;	DB 8;	Length 675;
Best Local Similarity	63.8%	Pred. No. 3.4e-60;		
Matches 387;	Conservative 0;	Mismatches 220;	Indels 0;	Gaps 0

Qy	400	AACAGAGTTT	GAGCAGAAAGCTTT	GTGCATTAATCT	CAGAGTCTGCTTTT	TTCGTAATTGGGT	459
Db	672	AAAAGAGCTTA	GACAGGAAGGTTT	GTGCACGTGCTAAT	CGGGGCTCTGTTT	CAATGATCATCTT	613
Qy	460	GGCCAACTTT	CAGCGAGATCGA	CCGAGGCTCGAATT	CTTGTGCTTTTGT	TCCGTTAGTGA	519
Db	612	GGCCCTGTTT	CAGCAATTGCA	CAGAGGACGGTATTTT	TGCGCCGGTGTG	CTCCGTTCTCTGA	553
Qy	520	ATGGCTTAAG	CTTGTTATTA	ACGAGATATTCATTT	CCCCAAATTCGAT	AGCTTAATCAAT	579

Accession	Sequence	Position
Db	552 ACTCCATGAGGCTTCTGATTATATGAGACCCGCTCTTCAACATGATGAGCTCGTAAAT	493
Oy	580 CCGTCACAAAGGAGGAGACAGAGAGTGTCTTAAAGTCCCTTGTCTACGTTTAC	639
Db	492 CAGTGCACGAGGAGGAAAAACAGAGGAAATGCTGAGAGGTCCACTATTATATGCTCGG	433
Oy	640 CTCTTCCTTTCTCTGCGGTTTCTTCTGGAAGAGTCTCCATCGGTATGATCTCGTAG	699
Db	432 TGCGTCGTTTAGCGTTTAACTCTTCTGCGCGAGATCCCCACGGAGACGCTCTCTTGT	373

OY	700	AAATGATCTGTGCGGGCGAATCGAATAGCTGATATPAAACGGACGTAACTTGTTGGCTCAACTA	759
Db	372	CGATGATAGCGGGTGGCCATGCTTTTGTGTGACATTTGTTGGAGAGCTATAGCTCAGCCA	313
OY	760	AGATACCTTCAACCCCAAGAAAGTTGGGGGAGGAGCATCTCCATGTTCACTCTTCGGCT	819
Db	312	AGTGCCACTTCAATCGGAAGAAGAGCTGGGGCGGGAGCATCTCGATGTTCACTTCTGGGTT	253
OY	820	TCTTCATCTTCATGCGAATTTACTTACTATTTACTCAAGCCTCTGGGGTACCTTCATCATGAAT	879

Db	252	TTCGCTCTCCGCGATGATGATGCTCTACTTCTCAAGCTGGGTACATGATGTTATATCT	193
Qy	880	GGGAAGACACCTTGCAGAGATAGCAATGATCTCAATGTGCCACGGATGCGATGCC	939
Db	192	GAGGAAGAGCACTTGGTAGCTGGGGCTGTGGCACTGACACGACAGTAGTGGAGTGGC	133
Qy	940	TACCATCAGCGATCAATTAGACGACAAATTTGCGTTCCTGTGCTACTATTTTAGCTG	999
Db	132	TTCTGTGATCCGAAAGTTTAGATGACACATATCTGTTCTCTTTGGCCACACAGTCTGTG	73

Qy	1000	CTTATTT	1006
Db	72	CTTTTCT	66

RESULT	14
CD405855/c	
LOCUS	658 bp mRNA linear EST 07-JUN-2003
CD405855	
DEFINITION	Gm_cK9937 Soybean induced by Salicylic Acid Glycine max cDNA 3'

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

CM nucleic - nucleic search, using sw model

Run on: March 16, 2006, 16:17:44 ; Search time 646 Seconds
(without alignments)
3002.043 Million cell updates/sec

Title: US-10-634-548-1
Perfect score: 1091
Sequence: 1 aaaaaagaataataacaaa.....ccgacctacaatgtttcc 1091

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Issued Patents NA: *
1: /cgn2_6/ptodata/1/ina/1 COMB.seq: *
2: /cgn2_6/ptodata/1/ina/5 COMB.seq: *
3: /cgn2_6/ptodata/1/ina/6A COMB.seq: *
4: /cgn2_6/ptodata/1/ina/6B COMB.seq: *
5: /cgn2_6/ptodata/1/ina/H COMB.seq: *
6: /cgn2_6/ptodata/1/ina/PC/US COMB.seq: *
7: /cgn2_6/ptodata/1/ina/PP COMB.seq: *
8: /cgn2_6/ptodata/1/ina/RE COMB.seq: *
9: /cgn2_6/ptodata/1/ina/backfile1.seq: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	67	6.1	7218	2	US-08-232-463-14
2	43.4	4.0	289	3	US-09-007-005-17
3	43.4	4.0	289	3	US-09-244-796-17
4	41.4	3.8	601	3	US-09-949-016-198946
5	41.4	3.8	767677	3	US-09-949-016-12147
6	41.4	3.8	767677	3	US-09-949-016-17361
7	40.6	3.7	143550	3	US-09-949-016-14143
8	39.4	3.6	524032	3	US-09-949-016-16928
9	39.4	3.6	524032	3	US-09-949-016-16929
10	39.4	3.6	524032	3	US-09-949-016-16930
11	39.4	3.6	524032	3	US-09-949-016-16931
12	39.4	3.6	529885	3	US-09-949-016-14340
13	39.4	3.6	529885	3	US-09-949-016-14341
14	39.4	3.6	529885	3	US-09-949-016-14342
15	39.4	3.6	529885	3	US-09-949-016-14343
16	39.4	3.6	529885	3	US-09-949-016-14344
17	39.4	3.6	529885	3	US-09-949-016-14345
18	39.4	3.6	529885	3	US-09-949-016-14346
19	39.4	3.6	529885	3	US-09-949-016-14347
20	39.4	3.6	264358	3	US-09-949-016-15725
21	38.8	3.6	72549	3	US-09-949-016-16477
22	38.2	3.5	395	3	US-09-894-844-45
23	37.4	3.4	832	3	US-09-621-976-2813
24	37	3.4	435	3	US-09-252-991A-6687

25	37	3.4	32010	3	US-09-949-016-13127	Sequence 13127, A
26	36.8	3.4	3117	3	US-09-614-221A-275	Sequence 275, App
27	36.4	3.3	76401	3	US-09-949-016-17153	Sequence 17153, A
28	36	3.3	87780	3	US-09-949-016-17011	Sequence 17011, A
29	36	3.3	90428	3	US-09-949-016-12564	Sequence 12564, A
30	36	3.3	152582	3	US-09-949-016-12086	Sequence 12086, A
31	36	3.3	152583	3	US-09-949-016-17390	Sequence 17390, A
32	36	3.3	152583	3	US-09-949-016-17391	Sequence 17391, A
33	35.8	3.3	288	3	US-09-313-294A-3968	Sequence 3968, App
34	35.8	3.3	601	3	US-09-328-111-530	Sequence 530, App
35	35.8	3.3	2769	3	US-09-949-016-4585	Sequence 4585, App
36	35.8	3.3	31573	3	US-09-949-016-16337	Sequence 16337, A
37	35.6	3.3	601	3	US-09-949-016-17919	Sequence 17919, A
38	35.6	3.3	601	3	US-09-949-016-17920	Sequence 17920, A
39	35.6	3.3	601	3	US-09-949-016-11809	Sequence 41809, A
40	35.6	3.3	601	3	US-09-949-016-41810	Sequence 41810, A
41	35.6	3.3	2586	3	US-09-248-796A-1910	Sequence 1910, App
42	35.6	3.3	46841	3	US-09-949-016-13466	Sequence 13466, A
43	35.6	3.3	69813	3	US-09-949-016-12455	Sequence 12455, A
44	35.6	3.3	69813	3	US-09-949-016-13905	Sequence 13905, A
45	35.6	3.3	69813	3	US-09-949-016-13906	Sequence 13906, A

ALIGNMENTS

RESULT 1
US-08-232-463-14
Sequence 14, Application US/08232463
Patent No. 5670367
GENERAL INFORMATION:
APPLICANT: DORNER, F.
APPLICANT: SCHRIFFLINGER, F.
APPLICANT: FALKNER, F. G.
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESSES:
ADDRESSER: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,463
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935,313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 IMMU
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: pTZgpc-Fls


```

1 PRIOR APPLICATION NUMBER: 60/237,768
2 PRIOR FILING DATE: 2000-10-03
3 PRIOR APPLICATION NUMBER: 60/231,498
4 PRIOR FILING DATE: 2000-09-08
5 NUMBER OF SEQ ID NOS: 207012
6 SOFTWARE: FASTSEQ for Windows Version 4.0
7 SEQ ID NO 198946
8 LENGTH: 601
9 TYPE: DNA
10 ORGANISM: Human
11 US-09-949-016-198946

```

Query Match	3.8%	Score 41.4;	DB 3;	Length 601;
Best Local Similarity	60.0%;	Pred. No. 0.019;		
Matches	69;	Conservative	0;	Mismatches 46;
			Indels	0;
			Gaps	0;

Qy	26	ATTTCCTTATCTTATTTGACTCTGCAAGATCTCTTCTCTCTCTCTCTCTCT	85
Dy	240	ATCATATCCATCTCTTTTACTTTGACAGATTTCTTATGTTCCATTTCTCAGACCTTCT	181
Qy	86	CCAACTCAGTTCCCTCCGCGTCATGCGACCACTTACTCTTATCTCCGATCAAT	140
Dy	180	CCCACTGACTCTCTCTCTCCATTTATTCACGACAGATGACTCTCTGATGAGAT	126

RESULT 5
US-09-949-016-12147

```

1  APPLICANT: KEMNER, J. Craig et al.
2  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
3  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
4  FILE REFERENCE: CL001307
5  CURRENT APPLICATION NUMBER: US/09/949,016
6  PRIOR FILING DATE: 2000-04-14
7  PRIOR APPLICATION NUMBER: 60/241,755
8  PRIOR FILING DATE: 2000-10-20
9  PRIOR APPLICATION NUMBER: 60/237,768
10 PRIOR FILING DATE: 2000-10-03
11 PRIOR APPLICATION NUMBER: 60/231,498
12 PRIOR FILING DATE: 2000-09-08
13 NUMBER OF SEQ ID NOS: 207012
14 SOFTWARE: FASTSEQ for Windows Version 4.0
15 SEQ ID NO 12147
16 LENGTH: 767677
17 TYPE: DNA
18 ORGANISM: Human
19 FEATURE:
20 NAME/KEY: misc_feature
21 LOCATION: (1)...(767677)
22 OTHER INFORMATION: n = A,T,C or G
23 US-09-949-016-12147

```

Query Match	3.8%	Score 41.4;	DB 3;	Length 767677;
Best Local Similarity	60.0%;	Pred. No. 1.2;		
Matches 69;	Conservative 0;	Mismatches 46;	Indels 0;	Gaps 0;

[illegible]

RESULT 6
US-09-949-016-17361
; Sequence 17361, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.

```

1  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
2  TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
3  FILE REFERENCE: CL001307
4  CURRENT APPLICATION NUMBER: US/09/949,016
5  CURRENT FILING DATE: 2000-04-14
6  PRIOR APPLICATION NUMBER: 60/241,755
7  PRIOR FILING DATE: 2000-10-20
8  PRIOR APPLICATION NUMBER: 60/237,768
9  PRIOR FILING DATE: 2000-10-03
10 PRIOR APPLICATION NUMBER: 60/231,498
11 PRIOR FILING DATE: 2000-09-08
12 NUMBER OF SEQ ID NOS: 207012
13 SOFTWARE: FastSeq for Windows Version 4.0
14 SEQ ID NO 17361
15 LENGTH: 767677
16 TYPE: DNA
17 ORGANISM: Human
18 FEATURE:
19 NAME/KEY: misc_feature
20 LOCATION: (1)...(767677)
21 OTHER INFORMATION: n = A,T,C or G
22 US-09-949-016-17361

```

Query Match	3.8%	Score 41.4;	DB 3;	Length 767677;
Best Local Similarity	60.0%;	Pred. NO. 1.2;		
Matches 69;	Conservative 0;	Mismatches 46;	Indels 0;	Gaps 0

Dy 26 ATTTCCTATACATAATGAATGTCAGAAATCCCTCTTCTCTTTCTCTCTCCTCCCT 85
Db 90697 ATCATATCCATCCTCTTACTTGACAGATTTCATGTGCATTTTCTCACACCTTCT 91
Dy 86 CCAACTGAGTCCCTCGCTGCATGGCAGCACTTACTCTATCTTCGATCAAT 140
Db 90757 CCCAATGAGCTCCCTCTGCCATTTATTCACGCAAGTGCATCTCTATAGT 90811

RESULT 7
US-09-949-016-14143

```

1  APPLICANT: VENTER, J. Craig et al.
2  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
3  TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
4  FILE REFERENCE: CLO01307
5  CURRENT APPLICATION NUMBER: US/09/949,016
6  CURRENT FILING DATE: 2000-04-14
7  PRIOR APPLICATION NUMBER: 60/241,755
8  PRIOR FILING DATE: 2000-10-20
9  PRIOR APPLICATION NUMBER: 60/237,768
10 PRIOR FILING DATE: 2000-10-03
11 PRIOR APPLICATION NUMBER: 60/231,498
12 PRIOR FILING DATE: 2000-09-08
13 NUMBER OF SEQ ID NOS: 207012
14 SOFTWARE: FastSeq for Windows Version 4.0.
15 SEQ ID NO 14143
16 LENGTH: 143550
17 TYPE: DNA
18 ORGANISM: Human
19 US-09-949-016-14143

```

Query Match	3.7%	Score 40.6;	DB 3;	Length 143550;
Best Local Similarity	49.3%;	Pred. No. 0.82;		
Matches 106;	Conservative 0;	Mismatches 109;	Indels 0;	Gaps 0;

[illegible]


```

RESULT 11
US-09-949-016-16931
Sequence 16931, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT FILING DATE: 2000-04-14
PRIORITY FILING DATE: 2000-10-20
PRIORITY FILING DATE: 2000-10-20
PRIORITY FILING DATE: 2000-10-03
PRIORITY FILING DATE: 2000-10-03
PRIORITY FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 16931
LENGTH: 524032
TYPE: DNA
ORGANISM: Human
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1) ... (524032)
OTHER INFORMATION: n = A,T,C or G
US-09-949-016-16931

```

```

RESULT 12
US-09-949-016-14340
; Sequence 14340, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949, 016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241, 755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237, 768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231, 498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14340
; LENGTH: 529885
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(529885)

```

```

US-09-949-016-14340
; OTHER INFORMATION: n = A,T,C or G

Query Match      3.6%; Score 39.4; DB 3; Length 52986;
Best Local Similarity 51.4%; Pred No. 4.1;
Matches 91; Conservative 0; Mismatches 86; Indels 0; Gaps 0;

QY    55 TTTCTTCTTCTTCTTCTTCTTCGTCCTCCGCCAAAGCAGTGCCCTCGGTCAATGGCAG 114
DB    180869 TTCTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1809228
QY    115 CACACTACTATCTGCAGATCAATCATCAGTTGTGTGCGTTCGGAAACAATTCTTGA 174
DB    180929 CTGCTCTAATTTTTCTTCTGCTTCAATCTTCGCTGTCTCTCTCTCTCTCTCTCT 1809886
QY    175 CGACGACACGGTTCGTTCGTCTCTCGAGCTCTTAATTTCTTCTCTCTGTTCATGAGTT 221
DB    180989 TTTCTTCTCTTCATTAATCATCTTCTCTCTCTTTTCTTCTTCTCTCTCTCTTCGATT 181045

```

```

RESULT 13
US-09-949-016-14341
; Sequence 14341, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO001307
; CURRENT APPLICATION NUMBER: US/09/949, 016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14341
; LENGTH: 529885
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..-(529885)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14341

```

Query Match		3.6%	Score 39.4	DB 3	Length 529865	
Best Local Similarity		51.4%	Pred No. 4.1			
Matches	91	Conservative	0	Mismatches	86	Indels 0; Gaps 0;
Dy	55	TTCTCTTCTTCTTCTTCTTCTTCCTCCCTCCTCAACTGATTCCTCGTGACAG 114				
Db	180869	TTCCTCTTCTCCTCCTCCTCCTCCTGCCTCTTCCTTCTTCCTTCTCTCCTCTCCT 180928				
Dy	115	CAACCTTACCCTCATCTCCGATCATCATCAATTGTGTGGTTCGGGAACAATCTTTGA 174				
Db	180929	CTGCTCATATTTTCTTCTCCTCTTCAATCTTCCTGTCCTCCCTCTTCTCTTCTTCTT 180988				
Dy	175	CGACTCACCGGTTCGTCTCTCCCTGACCTTCCTGATTTCTCTCCTGTTTCATTTGTT 231				
Db	180989	TTCTCTCTCTTCATATCATCTTCCTCTTCTTTTCTTCTCTCTCCTCCTTCGATT 181045				

RESULT 14
US-09-949-016-14342
; Sequence 14342, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 16, 2006, 18:40:51 ; Search time 3417 Seconds
(without alignments)
2640.296 Million cell updates/sec

Title: US-10-634-548-1

Perfect score: 1091
Sequence: 1 aaaaaagaataataacaa.....ccgaacttaaatgtttcc 1091

Scoring table: IDENTITY_NUC
Gap 10.0 , Gapext 1.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Published Applications NA Main:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 3: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
- 4: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
- 6: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
- 8: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1091	100.0	1091	US-10-634-548-1	Sequence 1, Appli
2	1015	93.0	1015	US-10-425-114-21028	Sequence 21028, A
3	977	89.6	980	US-10-425-114-14606	Sequence 14606, A
4	678.6	62.2	1062	US-10-425-114-29129	Sequence 29129, A
5	404	37.0	1897	US-10-634-548-3	Sequence 3, Appli
6	337	30.9	1007	US-10-425-114-20410	Sequence 20410, A
7	284.6	27.0	1033	US-10-425-114-26593	Sequence 26593, A
8	294.6	27.0	1085	US-10-425-114-22326	Sequence 22326, A
9	294.6	27.0	1974	US-10-437-963-77300	Sequence 77300, A
10	294.6	27.0	2514	US-10-437-963-77301	Sequence 77301, A
11	289	26.5	798	US-10-425-114-23255	Sequence 23255, A
12	289	26.5	1123	US-10-425-115-118753	Sequence 118753, A
13	278.8	25.6	976	US-10-425-114-26551	Sequence 26551, A
14	275.4	25.2	937	US-10-425-114-26454	Sequence 26454, A
15	274.8	25.2	723	US-10-425-114-25671	Sequence 25671, A
16	274.8	25.2	753	US-10-425-114-26498	Sequence 26498, A
17	270.6	21.1	674	US-10-425-114-16877	Sequence 16877, A
18	187.4	17.2	516	US-10-424-599-54356	Sequence 54356, A
19	171	15.7	1091	US-10-425-114-26298	Sequence 26298, A
20	169	15.5	648	US-10-425-114-16929	Sequence 16929, A
21	169	15.5	666	US-10-425-114-4925	Sequence 4925, Ap
22	169	15.5	955	US-10-425-115-118754	Sequence 118754, A
23	167.2	15.3	688	US-10-425-114-22387	Sequence 22387, A

24	167.2	15.3	907	US-10-767-701-11210	Sequence 11210, A
25	164.6	15.1	545	US-10-425-114-20484	Sequence 20484, A
26	161.6	14.8	1046	US-10-425-114-19285	Sequence 19285, A
27	161.6	14.8	1430	US-10-424-599-71025	Sequence 71025, A
28	160	14.7	1241	US-10-739-930-4866	Sequence 4866, Ap
29	155.8	14.3	795	US-10-425-114-13412	Sequence 13412, A
30	155.4	14.2	1225	US-10-425-114-26462	Sequence 26462, A
31	155.4	14.2	1285	US-10-425-115-172503	Sequence 172503, A
32	152.8	14.0	876	US-10-437-963-97687	Sequence 97687, A
33	147.2	13.5	944	US-10-425-114-19979	Sequence 19979, A
34	146.8	13.5	992	US-10-424-599-11249	Sequence 11249, A
35	146.4	13.4	699	US-10-425-114-25451	Sequence 25451, A
36	145.4	13.3	867	US-10-425-114-25373	Sequence 25373, A
37	144	13.2	948	US-10-437-963-38063	Sequence 38063, A
38	143	13.1	818	US-10-437-963-38063	Sequence 38063, A
39	143	13.1	1029	US-10-425-114-19839	Sequence 14839, A
40	141.4	13.0	1193	US-10-739-930-1297	Sequence 1297, Ap
41	136.4	12.5	1162	US-10-634-548-5	Sequence 5, Appli
42	131.2	12.0	1281	US-10-767-701-11596	Sequence 11596, A
43	127.8	11.7	1106	US-10-425-114-28780	Sequence 28780, A
44	126	11.5	1044	US-10-425-114-26552	Sequence 26552, A
45	116.6	10.7	1057	US-10-767-701-11766	Sequence 11766, A

ALIGNMENTS

RESULT 1
US-10-634-548-1
Sequence 1, Application US/10634548
Publication No. US20040045051A1
GENERAL INFORMATION:
APPLICANT: No. US20040045051A1r1s, Susan R
APPLICANT: Lincoln, Kim
APPLICANT: Abad, Mark Scott
APPLICANT: Elfers, Robert
APPLICANT: Hartshyker, Karen Kindle
APPLICANT: Hirschberg, Joseph
APPLICANT: Karunamadaa, Balasubramanian
APPLICANT: Moshiri, Farhad
APPLICANT: Stein, Joshua C.
APPLICANT: Valentini, Henry E.
APPLICANT: Venkatesh, Tyamagondlu V.
TITLE OR INVENTION: Tocopherol biosynthesis related genes and used thereof
FILE REFERENCE: Ren-01-135
CURRENT APPLICATION NUMBER: US/10/634,548
CURRENT FILING DATE: 2003-08-05
PRIOR APPLICATION NUMBER: us 60/400,689
PRIOR FILING DATE: 2002-08-05
NUMBER OF SEQ ID NOS: 79
SOFTWARE: PatentIn version 3.2
SEQ ID NO 1
LENGTH: 1091
TYPE: DNA
ORGANISM: Arabidopsis thaliana
US-10-634-548-1

Query Match 100.0%; Score 1091; DB 7; Length 1091;
Best Local Similarity 100.0%; Pred. No. 5.5e-312;
Matches 1091, Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAAAAGATTAATTAACAAATATCAATTCCTATATGACTGTCAAGATTCCT 60
DB 1 AAAAAAGATTAATTAACAAATATCAATTCCTATATGACTGTCAAGATTCCT 60
QY TCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 120
DB TCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 120
QY TACCTATATCCGATCAATCAATGTTGTCGTTGCGGACAACTCTTTGACGACTC 180
DB TACCTATATCCGATCAATCAATGTTGTCGTTGCGGACAACTCTTTGACGACTC 180

```

QY 181 ACCGGTCTGTTCTCTGGGCTTCTGATTTCTCTCTGTTGATTTGATGACCGGAA 240
DB 181 ACCGGTCTGTTCTCTGGGCTTCTGATTTCTCTCTGTTGATTTGATGACCGGAA 240
QY 241 TGGGCTCTGCTACGATTAAGTGTCTGTCTGTATCTTTGACGAGTTCGACGA 300
DB 241 TGGGCTCTGCTACGATTAAGTGTCTGTCTGTATCTTTGACGAGTTCGACGA 300
QY 301 ATTGGCTGTCATGACGTGGAGCAACCGGAGAGTCTGTGTGAGCATACGCGCTTG 360
DB 301 ATTGGCTGTCATGACGTGGAGCAACCGGAGAGTCTGTGTGAGCATACGCGCTTG 360
QY 361 TCTTAAGCTTGGAGAGTCTCAACGAGAAAGTCAATCAACAGATTGACGAGAAAGC 420
DB 361 TCTTAAGCTTGGAGAGTCTCAACGAGAAAGTCAATCAACAGATTGACGAGAAAGC 420
QY 421 TTGTGCTATATCTCTGAGGTCTGTTTGTGATCTTGCTGCTGCGCAATCTTGCGGATCGA 480
DB 421 TTGTGCTATATCTCTGAGGTCTGTTTGTGATCTTGCTGCTGCGCAATCTTGCGGATCGA 480
QY 481 CCGAGGCTCGATCTGTTGCTGTTTGTGATCTTGCTGCTGCGCAATCTTGCGGATCGA 540
DB 481 CCGAGGCTCGATCTGTTGCTGTTTGTGATCTTGCTGCTGCGCAATCTTGCGGATCGA 540
QY 541 ACGGACTATCCATTTCCCAATTCGATGCTATCAATCCGTCACAGAGAGGAGAG 600
DB 541 ACGGACTATCCATTTCCCAATTCGATGCTATCAATCCGTCACAGAGAGGAGAG 600
QY 601 CAGAAAGTTCCTTAAGGTCTTGTGCTGATCTGATCTGCTTCTTCTGCGGTTT 660
DB 601 CAGAAAGTTCCTTAAGGTCTTGTGCTGATCTGATCTGCTTCTTCTGCGGTTT 660
QY 661 TCTTCTGAGAGAGTCTCTATCCGATGATCTGCTGAGCATGATGTGTGCGAGT 720
DB 661 TCTTCTGAGAGAGTCTCTATCCGATGATCTGCTGAGCATGATGTGTGCGAGT 720
QY 721 GAATAGCTGATATATAGGAGCGTAAAGTTGGGTCAATAGATACCTTACCAACCGAAGAA 780
DB 721 GAATAGCTGATATATAGGAGCGTAAAGTTGGGTCAATAGATACCTTACCAACCGAAGAA 780
QY 781 AGAGTTGGGCGAGAAAGCATCTTCAGATGATCTTCTGAGTCTTCAATCCGATTAAC 840
DB 781 AGAGTTGGGCGAGAAAGCATCTTCAGATGATCTTCTGAGTCTTCAATCCGATTAAC 840
QY 841 TTTACTATTACCAAGCCTTGGGTACCTTCAATGAACTGGGAAACGACCTTGCAGAGAG 900
DB 841 TTTACTATTACCAAGCCTTGGGTACCTTCAATGAACTGGGAAACGACCTTGCAGAGAG 900
QY 901 TAGCAATGCTCAATGCTGCGCAGGATGCTGAGTCCGCTAACCCATCAACGATCAATTAG 960
DB 901 TAGCAATGCTCAATGCTGCGCAGGATGCTGAGTCCGCTAACCCATCAACGATCAATTAG 960
QY 961 ACGACAAATTTTGGGTCTCTGCGCTACTATTTTAAAGTCTTAAAGTTTGGATATT 1020
DB 961 ACGACAAATTTTGGGTCTCTGCGCTACTATTTTAAAGTCTTAAAGTTTGGATATT 1020
QY 1021 AGATTATCCCTCATAAACCGAATGTGATATACGATATTTTATGAAATCCGACTTA 1080
DB 1021 AGATTATCCCTCATAAACCGAATGTGATATACGATATTTTATGAAATCCGACTTA 1080
QY 1081 CAATGTTTCC 1091
DB 1081 CAATGTTTCC 1091

```

```

RESULT 2
US-10-425-114-21028
; Sequence 21028, Application US/10425114
; Publication No. US2004003488A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.

```

```

; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(5313)B
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 21028
; LENGTH: 1015
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana columbia
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3278-077-D5_FLI
US-10-425-114-21028

```

```

Query Match 93.0%; Score 1015; DB 7; Length 1015;
Best Local Similarity 100.0%; Pred. No. 1,7e-289; Mismatches 0; Gaps 0;
Matches 1015; Conservative 0; Indels 0;

```

```

QY 74 CTTGCTCTCTCCCAACCTGAGTCCCTCCGTCATGAGCAACCTTACCTATCTCC 133
DB 1 CTTGCTCTCTCCCAACCTGAGTCCCTCCGTCATGAGCAACCTTACCTATCTCC 60
QY 134 GATCAATCATCATGTTGTGTGCTTGGGAAACACTTTTGAAGTCAACCGGTTCTGTTTC 193
DB 61 GATCAATCATCATGTTGTGTGCTTGGGAAACACTTTTGAAGTCAACCGGTTCTGTTTC 120
QY 194 TCTGCGCTCTGATTTCTTCTCTGTTTCAATGTTTGAACCGGAATGGCTCTGCTAC 253
DB 121 TCTGCGCTCTGATTTCTTCTCTGTTTCAATGTTTGAACCGGAATGGCTCTGCTAC 180
QY 254 TCAGTTAGTGTGCTGCTGCTCTGATCTCTTCAAGATTCGAGATTCGAGATTCGAGT 313
DB 181 TCAGTTAGTGTGCTGCTGCTCTGATCTCTTCAAGATTCGAGATTCGAGATTCGAGT 240
QY 314 TGAAGTGGAGCAACCGTGGAGTGTGAGATGAGCATACGCGTTGTCTTAAGCTTGA 373
DB 241 TGAAGTGGAGCAACCGTGGAGTGTGAGATGAGCATACGCGTTGTCTTAAGCTTGA 300
QY 374 GAGTCTCAACGAGAAACGATTCACAGAGTTTGAAGCAAGAAAGTGTGATATCT 433
DB 301 GAGTCTCAACGAGAAACGATTCACAGAGTTTGAAGCAAGAAAGTGTGATATCT 360
QY 434 CTCAGGCTGCTTTCGATCTTGGTGGGCAATCTTCAAGCGATTCAGCGAGCTCGATA 493
DB 361 CTCAGGCTGCTTTCGATCTTGGTGGGCAATCTTCAAGCGATTCAGCGAGCTCGATA 420
QY 494 CTTTGTGCTTTTGTTCGTTAGTGAATGCTTAAGCTTGTATTAACGATATCCAT 553
DB 421 CTTTGTGCTTTTGTTCGTTAGTGAATGCTTAAGCTTGTATTAACGATATCCAT 480
QY 554 TTCCCAATTCGATGCTTAACCAATCCGTCACAGAGAAAGGAGAGAGAGTGTCT 613
DB 481 TTCCCAATTCGATGCTTAACCAATCCGTCACAGAGAAAGGAGAGAGTGTCT 540
QY 614 TAAAGTCTTGTGTTCAAGTCTTCTGAGCTCTTCTTCTGCGGTTTCTTCTGAGAGA 673
DB 541 TAAAGTCTTGTGTTCAAGTCTTCTGAGCTCTTCTTCTGCGGTTTCTTCTGAGAGA 600
QY 674 GTCTCTATCGGTATGATCTGTTAGCAATGATGTGTGCGAGTGAATAGTGATAT 733
DB 601 GTCTCTATCGGTATGATCTGTTAGCAATGATGTGTGCGAGTGAATAGTGATAT 660
QY 734 AATGGAAGTAAAGTTGGGTCAACTAAGATCTTCAACCAACCAAGAAAGTGTGGCAG 793
DB 661 AATGGAAGTAAAGTTGGGTCAACTAAGATCTTCAACCAACCAAGAAAGTGTGGCAG 720
QY 794 AAGCATCTCATGTTCAATCTTGGGCTTCTTCAATCTCATGCAATCTTACTATTAATCTC 853
DB 721 AAGCATCTCATGTTCAATCTTGGGCTTCTTCAATCTCATGCAATCTTACTATTAATCTC 780

```

QY	854	AAGCCTTGAGTACCTTCACATGAACCTGGAAAAGACCTTCGACAGAGTGTGCAATGGCTTC	913
Db	781	AAGCCTTGAGTACCTTCACATGAACCTGGAAAAGACCTTCGACAGAGTGTGCAATGGCTTC	840
QY	914	AATGTCGCACCGTAGTCGAGTCGCTACCCATCAACGATCAATTAGAAGACATATTTTC	973
Db	841	AATGTCGCACCGTAGTCGAGTCGCTACCCATCAACGATCAATTAGAAGACATATTTTC	900
QY	974	GGTTCCTCTGGCTACTATTATTTAGCTGCTTATTTAAGTTTCGGATATTAAGTTATATCCCTTC	103
Db	901	GGTTCCTCTGGCTACTATTATTTAGCTGCTTATTTAAGTTTCGGATATTAAGTTATATCCCTTC	960
QY	1034	ATPAACCGAATGTGTATATACGATATTTTATATGAATTCGACCTTACAAAGTTC	1088
Db	961	ATPAACCGAATGTGTATATACGATATTTTATATGAATTCGACCTTACAAAGTTC	1015

```

RESULT 3
US-10-425-114-14606
Sequence 14606, Application US/10425114
Publication No. US20040034588A1
GENERAL INFORMATION:
APPLICANT: Liu, Jingdong
APPLICANT: Zhou, Yihua
APPLICANT: Kovalic, David K.
APPLICANT: Screen, Steven E.
APPLICANT: Tabaska, Jack E
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(53113)B
CURRENT APPLICATION NUMBER: US/10/425,114
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 73128
SEQ ID NO 14606
LENGTH: 980
TYPE: DNA
ORGANISM: Arabidopsis thaliana
FEATURE:
OTHER INFORMATION: Clone ID: LIB23-048-A2_FLI
US-10-425-114-14606

```

	Query Match	89.6%	Score 977	DB 7	Length 980
	Best Local Similarity	100.0%	Pred. No. 3.1e-278		
	Matches 977	Conservative	0	Mismatches	0
				Indels	0
				Gaps	0
Qy	115	CAACCTTACCTTATCTCCGATCAATATCAATCACTTGTGTGGTTCGGAACTCTTTGA	174		
Db	1	CAACCTTACCTTATCTCCGATCAATATCAATCACTTGTGTGGTTCGGAACTCTTTGA	60		
Qy	175	CGACTCACCGGTTCTGTTTCTCCTGGCTTTGATTTCTTCTCCTTGTTCATTGGTTTGA	234		
Db	61	CGACTCACCGGTTCTGTTTCTCCTGGCTTTGATTTCTTCTCCTTGTTCATTGGTTTGA	120		
Qy	235	CCGGAATGGGCTCTGCTACTCAGTTTACGNGCTCGTGGTCTCGATCTCTTGACAGCTTG	294		
Db	121	CCGGAATGGGCTCTGCTACTCAGTTTACGNGCTCGTGGTCTCTGATCTCTTGACAGCTTG	180		
Qy	295	CGACGAATTCGCTGTGTGCATGAACGTCGAGCAACCGTGGCAGTGTGGTGTGAGCATACG	354		
Db	181	CGACGAATTCGCTGTGTGCATGAACGTCGAGCAACCGTGGCAGTGTGGTGTGAGCATACG	240		
Qy	355	CGCTTGTCTTAACTTCGAGAGTCTCACCAACGGAAACGTCAATGCAACAGATTTGAGA	414		
Db	241	CGCTTGTCTTAACTTCGAGAGTCTCACCAACGGAAACGTCAATGCAACAGATTTGAGA	300		
Qy	415	GAAGACTTGTGCATATTAATCTCAGAGTCTGCTTTTGTACTTGGTGAGCAATCTTCAGCG	474		
Db	301	GAAGACTTGTGCATATTAATCTCAGAGTCTGCTTTTGTACTTGGTGAGCAATCTTCAGCG	360		
Qy	475	GATGCAACCGAGGCTGATATCTTGTGCTTTTGTTCCTGATTAAGATATGCTTAAGGCTTG	534		
Db	361	GATGCAACCGAGGCTGATATCTTGTGCTTTTGTTCCTGATTAAGATATGCTTAAGGCTTG	420		

OY	535	TTATTAAGGACTATCCATTTTCCCAATTGCAATGCTATCAAAATCCGCACAGAAG	594
Db	421	TTATTAAGGACTATCCATTTTCCCAATTGCTATCAAAATCCGCACAGAAG	480
OY	595	GGAGAGCAGAAGATTGCTTAAAGGTCCTTGTTCTAGCTTCTCTCTCTG	654
Db	481	GGAGAGCAGAAGATTGCTTAAAGGTCCTTGTTCTAGCTTCTCTCTCTG	540
OY	655	CGGTTTTCTTCTGAGAGAGTCTCCTATCGGTATGATCTCGTTAGCAATGATGTGGTG	714
Db	541	CGGTTTTCTTCTGAGAGAGTCTCCTATCGGTATGATCTCGTTAGCAATGATGTGGTG	600
OY	715	GCGATGGAATGCTGATTAATATGGAAGTAAGTTGGGCACTAAGATACCTTAACAAC	774
Db	601	GCGATGGAATGCTGATTAATATGGAAGTAAGTTGGGCACTAAGATACCTTAACAAC	660
OY	775	CAAGAAAGATTGGGCGAGAAGCATCTCCATGTTTCATCTCGGCTCTTCATCTCATG	834
Db	661	CAAGAAAGATTGGGCGAGAAGCATCTCCATGTTTCATCTCGGCTCTTCATCTCATG	720
OY	835	CATTACTTACTATTACTCAAGCCTTGGGATCCTTCACTGAACCTGGAAAGACCTTGC	894
Db	721	CATTACTTACTATTACTCAAGCCTTGGGATCCTTCACTGAACCTGGAAAGACCTTGC	780
OY	895	AGAGAGTGCATAGTCTCAATGTCGCGCAAGTGTGAGAGTCGTAACCATCAACGATC	954
Db	781	AGAGAGTGCATAGTCTCAATGTCGCGCAAGTGTGAGAGTCGTAACCATCAACGATC	840
OY	955	AATTAGACGAATATTTTCGTTCTCTGCGTACTATTTTAGCTGTTATTAAGTTTCG	1014
Db	841	AATTAGACGAATATTTTCGTTCTCTGCGTACTATTTTAGCTGTTATTAAGTTTCG	900
OY	1015	GATATTAGATTAAATCCCTCATPAAACGAATGTGATATACGATTTTTTTTAATGAATCCG	1074
Db	901	GATATTAGATTAAATCCCTCATPAAACGAATGTGATATACGATTTTTTTTAATGAATCCG	960
OY	1075	ACCTTAGAAATGTTCC 1091	
Db	961	ACCTTAGAAATGTTCC 977	

RESULT 4
 US-10-425-114-29129
 Sequence 29129, Application US/10425114
 Publication No. US2004003488A1
 GENERAL INFORMATION:
 APPLICANT: Liu, Jindong
 APPLICANT: Zhou, Yihua
 APPLICANT: Kovacic, David K.
 APPLICANT: Screen, Steven E
 APPLICANT: Tadaska, Jack E
 APPLICANT: Cao, Yongwei
 TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
 TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
 FILE REFERENCE: 38-21(5313)B
 CURRENT APPLICATION NUMBER: US/10/425,114
 CURRENT FILING DATE: 2003-04-28
 NUMBER OF SEQ ID NOS: 73128
 SEQ ID NO 29129
 LENGTH: 1062
 TYPE: DNA
 ORGANISM: Brassica napus
 FEATURE:
 OTHER INFORMATION: Clone ID: LIB82-020-B9_FLI
 US-10-425-114-29129
 Query Match 62.2% Score 678.6; DB 7; Length 1062;
 Best Local Similarity 82.6% Pred. No. 8,1e-190;
 Matches 777; Conservative 0; Mismatches 164; Indels 0; Gaps 0;

```

Db      46 ATGCGGCGAGCGCTTACCTTTATCTCCGGGTAGCCATCAGTTATGCGGATTAAGCAACAG 105
Qy      168 TCTTTAGCACTCAACCGGTTCTGTTCTCCGGCTTCTGATTTCTTCTGTTTCTT 227
Db      106 TTTGGTATTAACGCTATACCTCCCGGTTCTGTTCTGCGAGTTTCTTCCCGCTTAACTC 165
Qy      228 GGTTTGACCGGAATGGGCTCTGCTACTCAGTTACGTCGTCGTTCTGATCTCTTCA 287
Db      166 GCGGTGAAGAAAGATCGGCTCGCGAGTCAAGTTAGGGGCTGCGAACCCCTGATCTTCA 225
Qy      288 GCAAGTTGACGAATTCGCTGTTGATACGTCGAGCCACCGTGGCAGTGGCTGTTGA 347
Db      226 GCGGCTTCGACTGATATCTATTCATGATACGTCGAGCCAGGTGGCAGTCTTAAGTGT 285
Qy      348 GCATACGCGCTTGTCTTAAGCTTGAGAGTCTCAACGAGCAACGATCAACAGAGT 407
Db      286 GCTTACGCGCTTGTCTTACTCTTTCGAGAGTCTCAACGAGCAACGATATTCACAGGA 345
Qy      408 TTGAGCAGAAAGCTTGTGCTATATCTCAAGGTCGCTTTTCTGATCTTGCSTGGCAATC 467
Db      346 TTGAGCAGAAAGCTTGTGCTATATCTCAAGGTCGCTTTTCTGAGCTCTCGGCGCAATC 405
Qy      468 TTCAGCGGATGACCGAGGCTCGATCTTGTGCTTTTGTTCGTTAGTGAATGCTTA 527
Db      406 TTCAGCGGATGACCGAGGCTCGATCTTGTGCTTTTGTTCGTTAGTGAATGCTTA 465
Qy      528 AGGCTTGTATTAACGGAATATTCATTCATTTCCCAATTCGATGATCAAAATCCGTCAC 587
Db      466 AGGCTTGTATTAACGGAATATTCATTCATTTCCCAATTCGATGATCAAAATCCGTCAC 525
Qy      588 AGAGAGGAGAGCAGAAAGAGTGTCTTAAGGTCCTTGTCTTCTGATCTTGTCTT 647
Db      526 AGGAGAGAGAGCAGAAAGAGTGTCTTAAGGTCCTTGTCTTCTGATCTTGTCTT 585
Qy      648 TTTCTGCGGTTTTTTCTGAGAGAGTCTCTATTCGGAATGATCTGTTAGCAATGATG 707
Db      586 GTTGTCTGCAATTTCTTCTGAGAGATTCCTTACCGGATGATATTCCTGCGCAATGATG 645
Qy      708 TGTGTGCGGATGGAATAGCTGATATTAATGGGAGCTAAGTTGGGTCATTAAGTAACT 767
Db      646 TGTGTGCGGATGGAATAGCTGATATTAATGGGAGCTAAGTAACTAAGTAACT 705
Qy      768 TACAACCCAGAAAGAGTGGGAGAGAGATCTTCATGTTCTATCTTGGCTTCTTCAATC 827
Db      706 TACAACCCAGAAAGAGTGGGAGAGAGATCTTCATGTTCTATTCGCTTCTTCAATC 765
Qy      828 TCCATCGCATTAATCTTAATCTTAATCAAGCTTGGGTAACCTTCAATGAATCGGAAACG 887
Db      766 TCCATCGCATTAATCTTAATCTTAATCAAGCTTGGGTAACCTTCAATGAATCGGAAACG 825
Qy      888 ACCTTGAGAGAGTGAATGCTCTCAATGGTGGCAGAGTAGTGAAGTGGCTAACCCATC 947
Db      826 ACCTTGAGAGAGTGGAGTGTCTCAATGGTGGCAGAGTAGTGAAGTGGCTAACCCATC 885
Qy      948 ACCGATCAATTAAGACGACAAATATTTCCGTTCTCTGCGTACTAATTTTAACTGCTTATTTA 1007
Db      886 ACCGATCAATTAAGACGACAAATATTTCCGTTCTCTGCGTACTAATTTTAACTGCTTATTTA 945
Qy      1008 AGTTTCGATTAATTAATTAATTCCTCTCAATTAACCGAATGTGT 1048
Db      946 AGTTTCGATTAATTAATTAATTCCTCTCAATTAACCGAATGTGT 986

```

```

RESULT 5
US-10-634-548-3
; Sequence 3, Application US/10634548
; Publication No. US20040045051A1
; GENERAL INFORMATION:
; APPLICANT: No. US20040045051A1file, Susan R
; APPLICANT: Lincoln, Kim
; APPLICANT: Abad, Mark Scott
; APPLICANT: Eilers, Robert
; APPLICANT: Hartsuyker, Karen Kindle

```

```

; APPLICANT: Hirsberg, Joseph
; APPLICANT: Karunanandaa, Balasubramanian
; APPLICANT: Moshiri, Farhad
; APPLICANT: Stein, Joshua C.
; APPLICANT: Valentin, Henry B.
; APPLICANT: Venkatesh, Tyamagondlu V.
; TITLE OF INVENTION: Tocopherol biosynthesis related genes and used thereof
; FILE REFERENCE: Ren-01-125
; CURRENT APPLICATION NUMBER: US/10/634,548
; CURRENT FILING DATE: 2003-08-05
; PRIOR APPLICATION NUMBER: us 60/400,689
; PRIOR FILING DATE: 2002-08-05
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO: 3
; LENGTH: 1897
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana

```

```

Query Match      37.0%; Score 404; DB 7; Length 1897;
Best Local Similarity 100.0%; Pred. No. 2,5e-108;
Matches 404; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 AAAAAAAAAAATTAATTAACAAATATCATTTCTTATCTTATTTGACTTGTCAAGATTCTCT 60
Db      1 AAAAAAAAAAATTAATTAACAAATATCATTTCTTATCTTATTTGACTTGTCAAGATTCTCT 60
Qy      61 TCTTCTTCTTCTTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 120
Db      61 TCTTCTTCTTCTTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 120
Qy      121 TACCTATCTCCGATCAATCATCATGTTGTGTCGAGTTCGAGAACATCTTTGACGATC 180
Db      121 TACCTATCTCCGATCAATCATCATGTTGTGTCGAGTTCGAGAACATCTTTGACGATC 180
Qy      181 ACCGGTCTGTTCTCTGCGCTTCTGATTTCTTCTCTGTTTCAATGTTGACCGGAA 240
Db      181 ACCGGTCTGTTCTCTGCGCTTCTGATTTCTTCTCTGTTTCAATGTTGACCGGAA 240
Qy      241 TGGGCTCTGCTACCTACCTACCTACCTACCTACCTACCTACCTACCTACCTACCTACCT 300
Db      241 TGGGCTCTGCTACCTACCTACCTACCTACCTACCTACCTACCTACCTACCTACCTACCT 300
Qy      301 ATTGCTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 360
Db      301 ATTGCTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 360
Qy      361 TCTTAAGCTTGAAGATCTCAACCAAGGAAAGTCAATCAACAG 404
Db      361 TCTTAAGCTTGAAGATCTCAACCAAGGAAAGTCAATCAACAG 404

```

```

RESULT 6
US-10-425-114-20410
; Sequence 20410, Application US/10425114
; Publication No. US2004003488A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Gao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 20410
; LENGTH: 1007
; TYPE: DNA

```


ORGANISM: *Gossypium hirsutum*
 FEATURE:
 OTHER INFORMATION: Clone ID: LIB3196-048-A11_FLI
 US-10-425-114-20410

Query Match 30.9%; Score 337; DB 7; Length 1007;
 Best Local Similarity 64.3%; Pred. No. 1.3e-88;
 Matches 505; Conservative 0; Mismatches 280; Indels 0; Gaps 0;

```

QY 280 TCTTTGAGCAGTTGCGAAGTTCGCTTGATGACGTGAGCCGCTGCGAGTGC 339
DB 181 TCTGAGCAGCCGCTGTAACCCCTCTATCTTCGAGATACCGCTGCTCCCTGCTCT 240
QY 340 TTGGTGAGCATAGCGCTGTTCTTTAGCTTCGAGCTTCACCAAGGAACGATTC 399
DB 241 TTGCTGGCGCTTAAGCTTCGCTTCGCTTCACCTTCGATTTCTACCTCAAAAGAGCTTC 300
QY 400 AACAGAGTTTGAGCAGAAAGCTTGATATCTCTAGCTCTGCTTTCTGACTTGGCT 459
DB 301 AGCAGATTTTAAGTAGAATAATGGTGATATATCTGATTACTTTTCCATTTCT 360
QY 460 GGGCAATCTTCAAGGATGACCAAGCTGATACTTTGCTTTGCTTCCGTTAGTGA 519
DB 361 GGGCAATTTTCAAGCAGCCGATGAGCTCGTACTTTCATCTGCTTCACCTTTCA 420
QY 520 ATGGCTTAAGGCTGTTATTAACGATATCCATTTCCCAATTCATGCTATCAAT 579
DB 421 ATGCTTAAGGCTGTTATTAATGATGCTCTCTTTGATGATGATCAAGCTTGAAT 480
QY 580 CCGTCAAGAGAGAGAGAGAGAGAGAGTCTTAAAGTCTTTGCTTACCTTCTAG 639
DB 481 CTGTTACTCGAGAGGAATCCCAAGAAATGCTTAAAGGAGCTTTGATTAATGTTGA 540
QY 640 CTCTCTTTCTCTGCGGTTTCTTCTGAGAGAGTCTCTATGCTATGATCTGTTAG 699
DB 541 TGTGATGATTAATGCTCTGCTGTTTGGCGTGAATCCCGGAGGATCTGCTG 600
QY 700 CAATGATGATGAGGAGAGAGAGAGAGAGTGAATGATTAATGAGAGAGAGTGA 759
DB 601 CAATGATGATGAGGAGAGAGAGAGAGAGAGTGAATTAATGAGAGAGAGAGTGA 660
QY 760 AGATACCTTACACCAAGAAAGAGTGGGAGAGAGATCTCATGTTCACTTCCGCT 819
DB 661 AGATTCCTTAATCAAGAGTGAAGTGGGAGAGATTTCCATGTTGTTCAAGAT 720
QY 820 TCTTCATCTCCATGCAATTAATTAATTAATCAAGCTTGGGTAACCTTCACTGA 879
DB 721 TCATCATTTCTAATGGAGTCTGATCTACTCACTGAGCTCTGGTTATTAATTA 780
QY 880 GGGAAAGCACTTGCAGAGATGACAAATGCTCAATGCTGCAAGCTGATCGAGTGC 939
DB 781 GGGGATATACCTCATATGAGGAGTCTTCAATTTCTAATGAGCACTGATTAATTC 840
QY 940 TACCATCATCGCATCAATTAAGCAGCAATATTCGTTCTCTGAGCTACTATTTAG 999
DB 841 TTCCAAATTCATGATGATTAATTAATTAATTTCTGTTCTCTGCTTCATGCTG 900
QY 1000 CTATTTAAGTTGAGATTAATTAATTAATTAATTAATTAATTAATTAATTA 1059
DB 901 CCAATTTAAGTTGAGATTAATTAATTAATTAATTAATTAATTAATTAATTA 960
QY 1060 TTTT 1064
DB 961 TTTT 965

```

RESULT 7
 US-10-425-114-26593
 Sequence 26593, Application US/10425114
 Publication No. US20040034888A1
 GENERAL INFORMATION:
 APPLICANT: Liu, Jindong
 APPLICANT: Zhou, Yihua

APPLICANT: Kovalic, David K.
 APPLICANT: Screen, Steven R
 APPLICANT: Tabaska, Jack E
 APPLICANT: Cao, Yongwei
 TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
 TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
 FILE REFERENCE: 38-21(5313)B
 CURRENT APPLICATION NUMBER: US/10/425,114
 CURRENT FILING DATE: 2003-04-28
 NUMBER OF SEQ ID NOS: 73128
 SEQ ID NO 26593
 LENGTH: 1033
 TYPE: DNA
 ORGANISM: *Oryza sativa*
 FEATURE:
 OTHER INFORMATION: Clone ID: LIB4371-003-F9_FLI
 US-10-425-114-26593

Query Match 27.0%; Score 294.6; DB 7; Length 1033;
 Best Local Similarity 62.9%; Pred. No. 4.6e-76;
 Matches 456; Conservative 0; Mismatches 269; Indels 0; Gaps 0;

```

QY 294 GCGAAGATTCGCTGTTGATGAGTGCAGCAAGCCGCTGCGAGTCTTGAGAGATAC 353
DB 192 GCGAGCCCGCGGCGCTGCGGAGCTGCGCGCCAGCTGCTCATCACCGGCGGCTAC 251
QY 354 GCGCTTGTCTTAAGCTTGCAGAGTCTCACCAAGCAAGCTCATTAACAGATTGAGC 413
DB 252 TCCCTGTGCGCGCTTGCAGAGGCTCACGCGCGGCGGCTCATCAGCAAGAACTGAGC 311
QY 414 AGAAGCTTGCATATCTCTAGCTCTGCTTTGCTTCTGCTTCTGCTGCTGCTGCTGCT 473
DB 312 AGAAGATTTGATAGTCTCTGAGAGTCTGCTTCTGCTTCTGCTTCTGCTTCTGCT 371
QY 474 GATGACGAGAGCTGATATCTTGTCTGCTTGTCTGCTTGTCTGCTTGTCTGCTT 533
DB 372 AATTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 431
QY 534 GTTATTAAGAGATATGATATTTCCCAATTCATGCTATCAATTCCTGCAAGAGAA 593
DB 432 CTGACTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 491
QY 594 GGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 653
DB 492 GGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 551
QY 654 GCGCTTCTTCTGAGAGAGTCTCTATGAGTGAATGCTGTTAGCAATGATGCTGCT 713
DB 552 GTTTGATCTTCTGCGCGAGTCTCTATGAGATGTTTCTGTTGATGATGATGATG 611
QY 714 GCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 773
DB 612 GGTATGCTTGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 671
QY 774 CCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 833
DB 672 GAGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 731
QY 834 GCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 893
DB 732 CTGATGCTGTTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 791
QY 894 CAGAGATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 953
DB 792 GGTAACTGCTCTGTTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 851
QY 954 CAATTAAGCAATATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1013
DB 852 GTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 911
QY 1014 GAGATA 1018
DB 912 GAGATA 916

```

RESULT 8

US-10-425-114-22326

; Sequence 22326, Application US/10425114
; Publication No. US2004003488A1

; GENERAL INFORMATION:

; APPLICANT: Liu, Jingdong

; APPLICANT: Zhou, Yihua

; APPLICANT: Kovalic, David K.

; APPLICANT: Screen, Steven E

; APPLICANT: Tadaka, Jack E

; APPLICANT: Cao, Yongwei

; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

; FILE REFERENCE: 38-21(5313)B

; CURRENT APPLICATION NUMBER: US/10/425,114

; FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 73128

; SEQ ID NO 22326

; LENGTH: 1085

; TYPE: DNA

; ORGANISM: Oryza sativa nipponbare

; FEATURE:

; OTHER INFORMATION: Clone ID: LIB3432-014-E9_FLI

US-10-425-114-22326

Query Match 27.0%; Score 294.6; DB 7; Length 1085;

Best Local Similarity 62.9%; Pred. No. 4.7e-76;

Matches 456; Conservative 0; Mismatches 269; Indels 0; Gaps 0;

294 GCGACGAATTCGCTGTCATGACGTCGAGCCACCGTGGACAGTGTCTTGTGAGCATAC 353

260 GCGAGCGCCGGGGGCTCGGAGACGCGCCGCAAGCTGTCATCACCGCGGGCGGTAC 319

354 GCGCTGCTTAAGCTTGAAGTTCACCAAGCAACGTCATTCACAGATTGAGC 413

320 TCCCTGTCGCGGCTTCGACGCGGCTCACGCGCGGCTCATCGAACAGAACCTGAGC 379

414 AGAAGCTTGTCATATCTCTCAGGCTGCTTTCGTAATGGTGGCAATCTTACG 473

380 AGAAAAATTGTCATGCTCTCCGAGTCTGTTCAATGCTTCTTGGCCACTATTCAGT 439

474 GGATCGACCGAGGCTGATCTTGTCTGTTGTTGCTTGAATGATGCTTAAAGGCTT 533

440 AATTGACAGAGCAGGTTCTTGGCGCAATGTCGGTGTGCTGAATCGATTAAGGCTT 499

534 GTTATTAACGACATATCCATTTCCCAATTGATGCTTAATCAATCCGTCACAGAGAA 593

500 CTGACCTACGGGCTCCGCTTCCACTGATGAAGCTTAAGTAAATCGGTGACCCGTGAA 559

594 GGGAGAGAGAGAGTGTCTTAAAGGCTCTTGTCTTGAATGCTTCTTCTTCT 653

560 GAAAAACAGAGAGATGCTGAGAGGCTCTGTAATGTCATGCTGCTGAGTCAAGT 619

654 GCGGTTTCTTCTGAGAGAGTCTCTATCGGTATGATCTCGTGAATGATGTGTGT 713

620 GTTTTGTCTTCTGCGGCAAGTCTCTATGAGGATGTTTGTGTGCTGATGATGAGTGT 679

714 GCGGATGAATGCTGATATTAATGGAGGTAAAGTTGGGTCACTTAAGTACTTCAAC 773

680 GGTATGCTTCTGTCGACATGTTGGAGAGGTATGCTCCGCAACCTGCAATTCAT 739

774 CCAAGAGAGAGTGGGAGAGAGATCTCAATGTCATCTTGGCTTCTTCAATCCATC 833

740 GAGAACAGAGCTGATAGAGAGATCTCAATGTCATCTTGGCTTCTTCAATCCATC 799

834 GCATCTTATCTATCTAAGCTTGGGTACCTTCAATGATGAGAGAGAGAGAGAGAG 893

800 CTGATGCTTCTAATCTTCTTGGGTACCTTCAATGATGAGAGAGAGAGAGAGAG 859

894 CAGAGATGAGAGATGCTCAATGCTGCGCAGAGATGAGAGAGAGAGAGAGAGAG 953

Db 860 GGTAACTGCTCTTGTGATAGACGCTACTGTAAGTGAAGTATCTCTGATGAT 919

Oy 954 CAATTAGACGACAAATATTCGTTCTCTGCTACTATTTTAACTTATTAAGTTTC 1013

Db 920 GTTGTGATGACAAATATCTGTTCTTGGCCACCAATGTTGGAGCTATCTGTATT 979

Oy 1014 GGATA 1018

Db 980 GGCTA 984

RESULT 9

US-10-437-963-77300

; Sequence 77300, Application US/10437963
; Publication No. US2004012343A1

; GENERAL INFORMATION:

; APPLICANT: La Rosa, Thomas J.

; APPLICANT: Kovalic, David K.

; APPLICANT: Zhou, Yihua

; APPLICANT: Cao, Yongwei

; APPLICANT: Mu, Wei

; APPLICANT: Boukharov, Andrey A.

; APPLICANT: Barbazuk, Brad

; APPLICANT: Li, Ping

; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With

; FILE REFERENCE: 38-21(53221)B

; CURRENT APPLICATION NUMBER: US/10/437,963

; FILING DATE: 2003-05-14

; NUMBER OF SEQ ID NOS: 204966

; SEQ ID NO 77300

; LENGTH: 1974

; TYPE: DNA

; ORGANISM: Oryza sativa

; FEATURE:

; OTHER INFORMATION: Clone ID: PAT_MRT4530_77211C.1

US-10-437-963-77300

Query Match 27.0%; Score 294.6; DB 7; Length 1974;

Best Local Similarity 62.9%; Pred. No. 6.5e-76;

Matches 456; Conservative 0; Mismatches 269; Indels 0; Gaps 0;

294 GCGACGAATTCGCTGTCATGACGTCGAGCCACCGTGGACAGTGTCTTGTGAGCATAC 353

1234 GCGAGCGCCGGGGGCTCGGAGACGCGCCGCAAGCTGTCATCACCGCGGGCGGTAC 1293

354 GCGCTGCTTAAGCTTGAAGTTCACCAAGCAACGTCATTCACAGATTGAGC 413

1294 TCCCTGTCGCGGCTTCGACGCGGCTCACGCGCGGCTCATCGAACAGAACCTGAGC 1353

414 AGAAGCTTGTCATATCTCTCAGGCTGCTTTCGTAATGCTTGGCGCAATCTTCAGC 473

1354 AGAAAAATTGTCATGCTCTCCGAGTCTGTTCAATGCTTCTTGGCCACTATTCAGT 1413

474 GGATCGACCGAGGCTGATCTTGTCTGTTGTTGCTTGAATGATGCTTAAAGGCTT 533

1414 AATTGACAGAGACGCTTCTGCGCAATGTCGGTGTGCTGCAATCGATTAAGGCTT 1473

534 GTTATTAACGACATATCCATTTCCCAATTGATGCTTAATCAATCCGTCACAGAGAA 593

1474 CTGACCTACGGGCTCCGCTTTCACATGATGAAGCTTAAGTAAATCGGTGACCCGTGAA 1533

594 GGGAGAGAGAGAGTGTCTTAAAGGCTCTTGTCTTGAATGCTTCTTCTTCT 653

1534 GGAAGACAGAGAGATGCTGAGAGGCTCTGTAATGTCATGCTGATGCTGCTGAGT 1593

654 GCGGTTTCTTCTGAGAGAGTCTCTATCGGTATGATCTGTTAGCAATGATGTGTGT 713

1594 GTTTTGTCTTCTGCGGCAATGCTCTAATGAGATGTTTGTGTGCTGATGATGAGTGT 1653

714 GCGATGAATGATGCTGATATTAATGGAGGTAAAGTTGGGTCACTTAAGTACTTCAAC 773

1654 GGTATGCTTCTGTCGACATGTTGGAGAGAGTATGAGTCTCCGCAACCTGCAATTCAT 1713

```

QY 774 CCAAGAAAGAGTTGGGAGAGAGATCTCATGTTCAATCTTGGCTTCTGATCCATC 833
DB 1714 GAGAACAGAGAGTGGATGAGAGATCTCAATGTTCAATTTCTGGCTTCTCTATGCA 1773
QY 834 GCATTACTTTACTTACTTACTCAAGCTTGGGATACCTTCACATGAATGGAGAACACCTTG 893
DB 1774 CTGATGCTGTTCTTCTCTCTGCTTGGTACTTCACTGCTGCTGGATCTGGCACTT 1833
QY 894 CAGAGATGAGATGCTCTCAATGCTGCCACGGTAGTGCAGTGCATACCGAT 953
DB 1834 GGTAACTGCTCTTGTGTGCAATGAGCTACTGATGAGAGTGAATTCCTGCAATGAT 1893
QY 954 CAATTAGACGACAAATATTTGGTTCCTGCTGCTACTATTTAGCTGTTATTAAGTTTC 1013
DB 1894 GTTGATGATGACAAATATCTCGTTCTTGGCACCAATGTTGGACGCTATCTGTAATTT 1953
QY 1014 GGATA 1018
DB 1954 GGCTA 1958

```

RESULT 10 US-10-437-963-77301

```

; Sequence 77301, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Mu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437, 963
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 77301
; LENGTH: 2514
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_77212C.1
US-10-437-963-77301

```

```

Query Match 27.0%; Score 294.6; DB 7; Length 2514;
Best Local Similarity 62.9%; Pred. No. 7.4e-76;
Matches 456; Conservative 0; Mismatches 269; Indels 0; Gaps 0;
QY 294 GCGAGCAATTCGCTGTCATGACGTGCGAGCCACCGGAGAGTGTGTGAGCATAC 353
DB 1672 GCGAGCGCCGCGGCTGCGGAGATCGCGCCGCAAGCTGCTCATACCGCGGCGCTAC 1731
QY 354 GCGCTTGTCTTAAGCTTGAAGTCTCAAGAGGAAAGTCAATTCACAGAGTTTGAAC 413
DB 1732 TCCGTGTCGCGGCTTCAAGCGGCTCAAGCGGCGGCTCATTCGAACAAACCTGAGC 1791
QY 414 AGAAGCTTGTGCAATATCTTCAAGTCTGCTTTTCTGACTTGGTGCGCAATCTTACG 473
DB 1792 AGAAAAATGTGCAATGTCTTCCGAGATCTGTTCATGTTCTTCTGGCCACTATTCAGT 1851
QY 474 GATCGACCGAGGCTGATCTTGTGCTTTGTTGCTGATGAAATGCTTAAGGCTT 533
DB 1852 AATTCGACAGAGACGCTTTCGCGGCAATGTCCGCTGTGCAACTGATTAAGGCTT 1911
QY 534 GTTATTTAAAGGATATTCATTTCCCAATTCGATGCTATCAATTCCTGCAAGAGAA 593
DB 1912 CTGACCTAAGGCTTCGCGCTTTCACATGATGAACTGATGAATGCGTGAACCGGAAA 1971

```

```

QY 594 GCGAGACGAGAAAGTTGCTTAAAGTCTTGTGTTCTACGTTCTAGCTCTTCTTCTCT 653
DB 1972 GAAAAACAGAGAAATTCGAGAGGCTCTCTGACTATGATGATGCTGCTGCTGACT 2031
QY 654 GCGGTTTTCTTGTGAGAGATCTCTATCGGATGATCTGTTAGCAATGATGTGCT 713
DB 2032 GTTTGGCTTTCGCGCCAGATCTCTATGAGATGTTTCGTTGATGATGATGATGAT 2091
QY 714 GCGCATGAAATAGCTGATATATATGAGACGTAATGTTGGTCACTAAGATACCTTAC 773
DB 2092 GGTATGCTTGTGTCATGTTGAGAGAGATGATGATGATGATGATGATGATGAT 2151
QY 774 CCAAGAAAGAGTTGGGAGAGAGATCTCCATGTTCAATCTTGGCTTCTTCAATCCATC 833
DB 2152 GAGAACAGAGCTGATGAGAGATCTCAATGTTCAATTTCTGCTTCTCTTATGCA 2211
QY 834 GCATTACTTTACTTACTTACTCAAGCTTGGGATACCTTCACATGAATGGAGAACACCTTG 893
DB 2212 CTGATGCTGTTCTTCTCTCTGCTTGGTACTTCACTGCTGCTGGATCTGGCACTT 2271
QY 894 CAGAGATGAGATGCTCTCAATGCTGCCACGGTAGTGCAGTGCATACCGAT 953
DB 2272 GGTAACTGCTCTTGTGTGCAATGAGCTACTGATGAGAGTGAATTCCTGCAATGAT 2331
QY 954 CAATTAGACGACAAATATTTGGTTCCTGCTGCTACTATTTAGCTGTTATTAAGTTTC 1013
DB 2332 GTTGATGATGACAAATATCTCGTTCTTGGCACCAATGTTGGACGCTATCTGTAATTT 2391
QY 1014 GGATA 1018
DB 2392 GGCTA 2396

```

RESULT 11 US-10-425-114-23255

```

; Sequence 23255, Application US/10425114
; Publication No. US20040034880A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jindong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53113)B
; CURRENT APPLICATION NUMBER: US/10/425, 114
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 23255
; LENGTH: 798
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB595-006-H9_FLI
US-10-425-114-23255

```

```

Query Match 26.5%; Score 289; DB 7; Length 798;
Best Local Similarity 62.6%; Pred. No. 1.8e-74;
Matches 451; Conservative 0; Mismatches 270; Indels 0; Gaps 0;
QY 286 CAGCAGTTGGAGCAATTCGCTGTCATGACGTGCGAGCAACGTTGAGCTTGGTG 345
DB 2 CGGCGCGGCGCGCCAGCGGTGCTGAGACGAGAGGTCATGTCATCAACCGCG 61
QY 346 GAGCATACGCGCTTGTCTTAAGCTTGAAGTCTCACCAAGCGAAAGTCAATCAAGA 405
DB 62 GCGCTACTCTCTTGTGCGGCTCTTCAACGAGCTCACCGAGCGGCTCATGAAAAGA 121
QY 406 GTTGAAGAGAAAGCTTGTGCAATATCTCAAGTCTGCTTTTGTGATCTTGGGCGAA 465
DB 122 GCTTGAAGAGAAAGTGTGCAAGTGTATCCGCGCTCTGTTCAATGATCTTGGGCC 181

```

```

Qy 466 TCTTCAGCGGATCGACGAGGCTGATACCTTGTGCTTTGTTCCGTAGTGAATGCT 525
Db 182 TGTTCAGCAATTCGACGAGGACGATTTTCGCGCGGTTGTCCGTCTCTGAATCTCA 241
Qy 526 TAAGGCTTGTATTAAAGGATATTCATTTCCCAATTGCGATCTAATCAATCCGTCA 585
Db 242 TGAAGCTTCTGATATATGAGTCTCGTCTTACACTGATGAAGCTCTGGTAAATCAATGA 301
Qy 586 CAGAGAGAGGAGGAGGAGAGAGTGTCTAAAGGCTCTTGTCTAGGTTCTAGCTCTTC 645
Db 302 CACGTGAAGGAAAACAGAGGAAATGCTGAGAGGTCCACTATATATGTCTGTGCTGC 361
Qy 646 TTTTCTGCGGTTTTCTTCTGAGAGAGTCTCTATCGATGATCTCGTTAGCAATGA 705
Db 362 TGTTCAGCGTTTATGCTCTTCTGCGGTGAGTCCCATCGGATGCTCTCTTGTGATGA 421
Qy 706 TGTGTGTGCGGATGAGATAGCTGATATATATGGAAGCTAATTTGGGTCAATGAATAC 765
Db 422 TGAAGCGGTGCGGATGATTTTGTCTGACATTTGTGGAGGAGGATAGGCTCAGCGAAGCTGC 481
Qy 766 CTTCACACCCAGAAAGAGTTGGGAGGAGATCTCATGTCATCTTGGCTTCTCA 825
Db 482 CATTCAATCGGAGAGAGAGCTGGGCGGAGACATCTGATGTCAATTTCTGTTTCTGC 541
Qy 826 TCTTCATCGCATTAATCTTACTATTACTCAAGCTTGGGTACCTTCAATGAATCGGAAA 885
Db 542 TGTCCGCGATGATGATGCTCTACTTCTCAAGCTGGGTACATTTGATTTATCTGGGAG 601
Qy 886 CGACCTTGACAGAGTAGCAATGCTCTCATATGTGCGACGATAGTGCATGCTGATCCA 945
Db 602 AGGCACTTGTAAAGCTGGCGCTGTGTGCACTAGCAGACAGATGAGAGTGCCTCTCG 661
Qy 946 TCACCGATCAATTAAGACAAATATTCGTTCCCTGCGCTACTATTTAGCTGCTATT 1005
Db 662 TGACCGAAGTTGTGATGATCAACATCTGTTCTTTGGCCACATGCTGATGCTTTTC 721
Qy 1006 T 1006
Db 722 T 722

```

```

RESULT 12
US-10-425-115-118753
; Sequence 118753, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 118753
; LENGTH: 1123
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_39790C.1
US-10-425-115-118753

```

```

Query Match 26.5%; Score 289; DB 8; Length 1123;
Best Local Similarity 62.6%; Pred. No. 2,2e-74;
Matches 451; Conservative 0; Mismatches 270; Indels 0; Gaps 0;

```

```

Qy 286 CAGCAGTTGCGACGAATTCGCTGTGCATGACGTGAGGACACGCTGGACAGTCTTGTG 345
Db 253 CGGCGCGCGCGCGCGCGCGGCTGTGCAAGACGCGGCGGCTGCTGCTCATCACCGCG 312

```

```

Qy 346 GAGCATACGCGCTTGTCTTAAGCTTGGAGAGTCTCACACAGCCGAACGTCATTCAACA 405
Db 313 GCGCTTACTCTTGTGCGCGGCTCTTCGACGAGCTCACCGGCGGCTCATCGAANAAGA 372
Qy 406 GTTGGACGAAGAGCTTGTGCAATATACCTCAGGTCTGCTTGTGATCTTGGGCGCA 465
Db 373 GCTTGAAGCAAGAGGTGTGCAAGTCTATCCGCGTCTGTTTCATCTTGGGCC 432
Qy 466 TCTTCAGCGGATCGACGAGGCTGATACCTTGTGCTTTTGTTCGTTAGTGAATGAGCT 525
Db 433 TGTTCAGCAATTCGACGAGGAGCAAGGATTTGCGCGGTTGTGTCGTTCTGAAATCCA 492
Qy 526 TAAGGCTTGTATTAAAGGATATTCATTTCCCAATTGCGATCTAATCAATCCGTCA 585
Db 493 TGAAGCTTCTGATATATGAGTCTCGTCTTCACTGATGAAGCTCTGGTAAATCAATGA 552
Qy 586 CAGAGAGGAGGAGGAGGAGAGTGTCTAAAGGCTCTTGTCTAGCTTCTAGCTCTTC 645
Db 553 CACGTGAAGGAAAACAGAGGAAATGCTGAGAGGTCCACTATATATGTCTGTGCTGC 612
Qy 646 TTTTCTGCGGTTTTCTTCTGAGAGAGTCTCTATCGATGATCTCGTTAGCAATGA 705
Db 613 TGTTCAGCGTTTATGCTCTTCTGCGGTGAGTCCCATCGGAGTGTCTCTTGTGATGA 672
Qy 706 TGTGTGTGCGGATGAGATAGCTGATATATATGGAAGCTAATTTGGGTCAATGAATAC 765
Db 673 TGAAGCGGTGCGGATGATTTTGTCTGACATTTGTGGAGAGGATAGGCTCAGCGAAGCTGC 732
Qy 766 CTTCACACCCAGAAAGAGTTGGGAGGAGACATCTCATGTCATCTTGGCTTCTTCA 825
Db 733 CATTCAATCGGAGAGAGAGCTGGGCGGAGACATCTGATGTTTCAATTTCTGTTTCTGC 792
Qy 826 TCTTCATCGCATTAATCTTACTATTACTCAAGCTTGGGTACCTTCAATGAATCGGAAA 885
Db 793 TGTCCGCGATGATGATGCTCTACTTCTCAAGCTGGGTACATTTGATTTATCTGGGAG 852
Qy 886 CGACCTTGACAGAGTAGCAATGCTCTCATATGTGCGACGATAGTGCATGCTGATCCA 945
Db 853 AGGCACTTGTAAAGCTGGCGCTGTGTGCACTAGCAGGACAGATGAGAGTGCCTCTCG 912
Qy 946 TCACCGATCAATTAAGACAAATATTCGTTCCCTGCGCTACTATTTAGCTGCTATT 1005
Db 913 TGACCGAAGTTGTGATGATCAACATCTGTTCTTTGGCCACATGCTGATGCTTTTC 972
Qy 1006 T 1006
Db 973 T 973

```

```

RESULT 13
US-10-425-114-26551
; Sequence 26551, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53113)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 26551
; LENGTH: 976
; TYPE: DNA
; ORGANISM: Trifolium aestivum
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB4312-010-C11_PL1
US-10-425-114-26551

```

Query Match 25.6%; Score 278.8; DB 7; Length 976;
Best Local Similarity 61.6%; Pred. No. 2.2e-71;
Matches 445; Conservative 0; Mismatches 277; Indels 0; Gaps 0;

294 GCGACGATTCGCTGTCATGACGTCGAGCCACCGTGGAGCTGCTTGGAGCATAC 353
Db GCGGCTCATGCTGCTGCGGAGCGGCGGCGACGCTGCTGACCGCGGCGCATAC 263
Qy GCGCTGCTTAAAGCTTGAAGTCTCACGAGGAAAGTCATTCACAGAGTTGAGC 413
Db TCCCTCGCGCGCTTGCAGCGCTCACCGAGCCCGCTGCTCAACAGAGCTTGGC 323
Qy AGAAGCTTGCATATCTCTCAGGCTGCTTTCGATCTTGGTGGCCCATCTTGGC 473
Db AGAAGCTTGCATATCTCTCAGGCTGCTTTCGATCTTGGTGGCCCATCTTGGC 383
Qy GATGACCGAGGCTCGATCTTGGCTGCTTTCGATCTTGGTGGCCCATCTTGGC 533
Db AATTCACGATGACGCTTCTTGGCAGCGGATGCTTCTGATCTTGGTGGCCCAT 443
Qy GTTATTAAGCATATTCATTTCCCAATTCATGATCTTAAATCCGTCACAGAGAA 593
Db CTCACCTACGCGGCTCGCTTCTACTCCGACGAGCTTAAATCCGTCACCGTGA 503
Qy GGGAGAGAGAAAGTGGCTTAAAGTCTTTCGATCTTGGTGGCCCATCTTGGC 653
Db GGAAGAGAGAAAGTGGCTTAAAGTCTTTCGATCTTGGTGGCCCATCTTGGC 563
Qy GCGCTTTCCTTCTGAGAGAGTCTCTATCCGATGATCTGATGAGCATGATGAT 713
Db GTTCTAGCTTTCGCGGAGCTCCCGCATCCGATGATCTTCTGATGAGCATGAT 623
Qy GGCATGAGATGCTGATATTAATGGAGCTTAAGTTGGCTCACTAAGATCTTAAAC 773
Db GGTACGCTTTCGCTGATGATGCTGAGAGAGTTCGCTGCTCAACCTCAATCAAC 683
Qy CCAAGAAAGATGGGCGAGAGAGCATCTCCATGCTTTCGCTGCTTTCATCTCCATC 833
Db AAGAAAGAGAGCTGGTGGGAGAGTCCGAGATGCTTTCGCTGCTTTCATCTCCATC 743
Qy GCATTAATCTTACTATCTCAAGCTTGGGATCTTCAATGATGAGAAACGACCTTG 893
Db CTGATGCTGCTTATCTTCTGCTGCTTGGTATCATGATGATGAGTGGAGTCAAGCGCTT 803
Qy CAGAGAGTGAAGATGCTCAATGCTGCGCACGAGTGCAGTGCATCCATCCGAT 953
Db GGTAAAGTGTCTGCTGCTGCGCACGCTGAGTGAATCTGTAATGAT 863
Qy CAATTAAGCAGATATTTGGTTCCTGCTGCTATTTTAACTGCTTATTAAGTTTC 1013
Db GTTGAAGTGAATATCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTTTC 923
Qy 1014 GG 1015
Db 924 GG 925

RESULT 14
US-10-425-114-26454
; Sequence 26454, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jindong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114

CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 26454
; LENGTH: 937
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB4185-014-E12_FLI
US-10-425-114-26454

Query Match 25.2%; Score 275.4; DB 7; Length 937;
Best Local Similarity 62.3%; Pred. No. 2.1e-70;
Matches 449; Conservative 0; Mismatches 271; Indels 1; Gaps 1;

286 CAGAGTTGCGACGAAATTCGCTGTCATGACGTCGAGCCACCGTGGCAATGCTGTG 345
Db CGGCGGCGCGCGCGCGCGAGGCTGCGAGAGCGGCGGCTGCTGCTGCTGCTGCTG 217
Qy GAGCATACGCGCTTGTAAAGCTTGAAGTCTCACCAAGGAAAGCTCATTCAGAGAA 405
Db GCGCTTACTCTTCTGCGCGCTTTCGACGAGCTCACCGAGCGGCGCTCATCGAGAA 277
Qy GTTGAAGAGAAAGCTTGTGATATCTTCAAGTCTGCTTTCGATCTTGGTGGCCCA 465
Db GCTTGAAGAGAGAGTGTGACGCTGCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTG 337
Qy TCTTCAAGGATTCAGAGGCTGATATCTTGTGCTTTCGCTTTCGCTTTCGATGATGCT 525
Db TGTGAGAGAAATTCAGAGAAAGAGTTCGCGCGGCTGCTGCTGCTGCTGCTGCTGCTG 397
Qy TAAAGCTTGTATTAAGAGATATCTTCCCAATTCATGATGCTTAAATCCGCA 585
Db TGAAGCTTGTATTAAGAGATATCTTCCGCTTCACTGATGAGCTGCTGATTAATCAG 457
Qy CAGAGAGAGAGAGAGAGAGAGTTCGCTTAAAGTCTTTCGCTTTCGCTTTCGCTTTC 645
Db CAGAGAGAGAGAGAGAGAGAGTTCGCTTAAAGTCTTTCGCTTTCGCTTTCGCTTTC 516
Qy TTTTCTGCGGCTTTCGCTGAGAGTCTCTATGCTGATGATCTTCTGATGATGATGAT 705
Db TGTTCAGGCTTTCGCTTTCGCTGAGAGTCTCTATGCTGATGATCTTCTGATGATGAT 576
Qy TGTGAGGAGAGAGAGAGATGATATTAATGAGAGAGTTCGCTTTCGCTTTCGCTTTC 765
Db TGAAGGAGAGAGAGAGTTCGCTTTCGCTTTCGCTTTCGCTTTCGCTTTCGCTTTC 636
Qy CTTAAGAGAGAGAGAGAGTTCGCTTTCGCTTTCGCTTTCGCTTTCGCTTTCGCTTTC 825
Db CATCAATCGAGAGAGAGAGTTCGCTTTCGCTTTCGCTTTCGCTTTCGCTTTCGCTTTC 696
Qy TCTTCATGCTTTCGCTTTCGCTTTCGCTTTCGCTTTCGCTTTCGCTTTCGCTTTCGCTTTC 885
Db TGTTCGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 756
Qy CGACCTGAGAGAGAGAGAGTTCGCTTTCGCTTTCGCTTTCGCTTTCGCTTTCGCTTTC 945
Db AGGACTTGTGAGAGAGAGAGTTCGCTTTCGCTTTCGCTTTCGCTTTCGCTTTCGCTTTC 816
Qy TCACGATCAATTAAGAGAGAGATTCGCTTTCGCTTTCGCTTTCGCTTTCGCTTTCGCTTTC 1005
Db TGACGAGAGTGTGAGAGAGAGATTCGCTTTCGCTTTCGCTTTCGCTTTCGCTTTCGCTTTC 876
Qy 1006 T 1006
Db 877 T 877

RESULT 15
US-10-425-114-25671
; Sequence 25671, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jindong

Job time : 3420 secs

APPLICANT: Zhou, Yihua
APPLICANT: Kovalic, David K.
APPLICANT: Screen, Steven E.
APPLICANT: Tabaska, Jack E.
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(5313)B
CURRENT APPLICATION NUMBER: US/10/425,114
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 73128
SEQ ID NO 25671
LENGTH: 723
TYPE: DNA
ORGANISM: Zea mays
FEATURE:
OTHER INFORMATION: Clone ID: LIB957-022-E3_FLI
US-10-425-114-25671

Query Match 25.2% Score 274.8; DB 7; Length 723;
Best Local Similarity 63.4%; Pred. No. 2.8e-70;
Matches 420; Conservative 0; Mismatches 242; Indels 0; Gaps 0;

QY 345 GGACATACGGCGCTTGTCTTAAGCTTCGAGAGTCTCACGAAAGCGAAAGTTCATTCAACAG 404
DB 7 GGCGCTACTCTCTTGTGCGCGTCTTCGACGAGCTCACGAGCGCGGCTCATCGAAAG 66
QY 405 AGTTGAGCAGAAAGCTTGTGCATATATCTCAGGTCTGCTTTTCTGTACTTGGCGTCCCA 464
DB 67 AGCTTGAGCAGAAAGTGTGACGTGCTGCGCGCTCTGTCATGTCATCTTGGCCC 126
QY 465 ATCTTCAGCGAGTGCACGAGGCTCGATATCTTGTGCTTGTTCGTTAGTGAATGCG 524
DB 127 CTGTTCAAGCATTTGACAGAAAGCAAGATTTCCGCGGTTGTCCTTCTGAACCTCC 186
QY 525 TTAAAGCTTGTATTAAAGCACTATCCATTTCCCAATTGATGCTAATCAAAATCCGTC 584
DB 187 ATGAGGCTTGTATATATGACTCGTCTACACTGATGAAGCTGTGTAAATCAGTG 246
QY 585 ACAAGAGAGGAGAGCAAGAGATGCTTAAAGTCTTGTCTTGAAGTCTTGAAGTCTT 644
DB 247 ACAGGTGAAGGAAACCAAGAAATGCTGAAGGTCCACTTATATGCTGGTGTGCTG 306
QY 645 CTTTCTCTGCGGTTTTCTTGTGAGAGAGTCTCCTATCGGTATGATCTCGTTAGCAATG 704
DB 307 CTGTTCAAGCTTTTATAGTCTTGTGCGGTAGTCCCATCGGATCGTCTCTTGTGATG 366
QY 705 ATGTGTGTGCGATGGAATAGCTGATTAATGGAAGTAAAGTTGGTCACTAAGATA 764
DB 367 ATGAGCGGTGCGATGTTTGTGCTGACATTTGTGGAGAGATGCTGCTCAGCGAAGCTG 426
QY 765 CCTTCAACCCAAAGAAAGTGGGCAAGAGCATCTTCATGTTCAATCTTGGGCTTCTC 824
DB 427 CCATTCAATCGAAGAAAGCTGGGCGGAGCACTGATGTTCAATTTCTGTTCTCTG 486
QY 825 ATCTCATCGCATTAATTATTAATCAAGCTTGGGTACCTTCAATGAACCTGGGAA 884
DB 487 CTGTCCGCGATGATGATCTCTTCTTCAAGCTTGGGTATCATTTGATGATGTTATCTGGAA 546
QY 885 ACGACCTTGCAAGAGTGAAGTGTCTCAATGTGTGCGCACGATGAGTGTGCTACCC 944
DB 547 GAGGCACTTGTGAAGCTGGCGCTTGTGCACTAGCAGGAGCAAGTGTGAGTGTGCTTCT 606
QY 945 ATCACCGATCAATTAGACGACATTTTGGTCTCTGCTGCTACTAATTTTGAAGCTTAT 1004
DB 607 GTGACCGAAGTGTATGATGACAAATATCTTCTTGTGCGCACCATGCTGGTGAAGCTTT 666
QY 1005 TT 1006
DB 667 CT 668

Search completed: March 16, 2006, 19:47:26

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 16, 2006, 18:15:06 ; Search time 475.Seconds
(without alignments)
5355.560 Million cell updates/sec

Title: US-10-634-548-1

Perfect score: 1091
Sequence: 1 aaaaaagaataataacaa.....ccgacctacaatgtttcc 1091

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 8023312 seqs, 1165852854 residues

Total number of hits satisfying chosen parameters: 16046624

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA New:*
1: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB_seq.*
2: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB_seq.*
3: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB_seq.*
4: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB_seq.*
5: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB_seq.*
6: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB_seq.*
7: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB_seq.*
8: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB_seq.*
9: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB_seq.*
10: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB_seq.*
11: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB_seq.*
12: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB_seq.*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB_seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	ID	Description
1	287.4	26.3	US-11-096-568A-19558	Sequence 19558, A
2	169	481	US-11-096-568A-14317	Sequence 14317, A
3	169	15.5	US-11-096-568A-1769	Sequence 1769, A
4	153.8	14.1	US-11-096-568A-14811	Sequence 14811, A
5	41	3.8	US-10-995-561-11236	Sequence 11236, A
6	40.8	3.7	US-11-136-527-7472	Sequence 7472, A
7	40.8	3.7	US-11-136-527-7376	Sequence 7376, A
8	39.8	3.6	US-09-925-065A-655075	Sequence 655075, A
9	39	3.6	US-09-925-065A-655075	Sequence 655075, A
10	38.6	3.5	US-09-925-065A-655076	Sequence 655076, A
11	38.2	3.5	US-11-143-401-45	Sequence 45, App1
12	37.8	3.5	US-09-925-065A-512623	Sequence 512623, A
13	37.8	3.5	US-10-330-773-832	Sequence 832, App
14	37.2	3.4	US-09-925-065A-718806	Sequence 718806, A
15	37.2	3.4	US-10-330-773-645	Sequence 645, App
16	36.6	3.4	US-09-925-065A-512622	Sequence 512622, A
17	36.4	3.3	US-10-330-773-156	Sequence 156, App
18	36.4	3.3	US-09-925-065A-693085	Sequence 693085, A
19	36.2	3.3	US-09-925-065A-693085	Sequence 693085, A
20	36.2	3.3	US-09-925-065A-693086	Sequence 693086, A

21	36	3.3	2148	12	US-11-136-527-313	Sequence 313, App
22	35.8	3.3	591	6	US-09-925-065A-501602	Sequence 501602, A
23	35.8	3.3	1908	7	US-10-932-182A-82752	Sequence 82752, A
24	35.8	3.3	1908	7	US-10-932-182A-82752	Sequence 82752, A
25	35.8	3.3	3096	12	US-11-127-817-6	Sequence 6, App1
26	35.8	3.3	189252	12	US-11-121-086-54	Sequence 54, App1
27	35.6	3.3	644	6	US-09-925-065A-795176	Sequence 795176, A
28	35.6	3.3	1653	6	US-09-925-065A-695563	Sequence 695563, A
29	35.6	3.3	1953	12	US-11-128-660-2	Sequence 2, App1
30	35.6	3.3	106610	7	US-10-330-773-585	Sequence 585, App
31	35.2	3.2	138808	7	US-10-330-773-905	Sequence 905, App
32	35	3.2	3399	9	US-11-096-568A-29372	Sequence 29372, A
33	35	3.2	25458	8	US-10-995-561-13390	Sequence 13390, A
34	35	3.2	321019	8	US-10-995-561-13204	Sequence 13204, A
35	35	3.2	611587	12	US-11-117-187-209	Sequence 209, App
36	34.8	3.2	540	6	US-09-925-065A-511072	Sequence 511072, A
37	34.8	3.2	516	6	US-09-925-065A-476732	Sequence 476732, A
38	34.8	3.2	645	6	US-09-925-065A-421329	Sequence 421329, A
39	34.8	3.2	646	6	US-09-925-065A-469081	Sequence 469081, A
40	34.8	3.2	742	6	US-09-925-065A-944264	Sequence 944264, A
41	34.8	3.2	1095	6	US-09-925-065A-673651	Sequence 673651, A
42	34.8	3.2	92969	7	US-10-330-773-178	Sequence 178, App
43	34.6	3.2	569	6	US-09-925-065A-284214	Sequence 284214, A
44	34.6	3.2	193789	12	US-11-112-908-55	Sequence 55, App1
45	34.6	3.2	215974	7	US-10-330-773-500	Sequence 500, App

ALIGNMENTS

RESULT 1	US-11-096-568A-19558	Application US/11096568A
Sequence 19558	US/11096568A	Publication No. US20060048240A1
GENERAL INFORMATION:		
APPLICANT: Alexandrov, Nickolai et al.		
TITLE OF INVENTION: Sequence-determined DNA Fragments and Corresponding Polypeptides		
TITLE OF INVENTION: Theby		
FILE REFERENCE: 2750-1592P052		
CURRENT APPLICATION NUMBER: US/11/096.568A		
CURRENT FILING DATE: 2005-04-01		
NUMBER OF SEQ ID NOS: 34471		
SEQ ID NO 19558		
LENGTH: 1198		
TYPE: DNA		
ORGANISM: Zea mays subsp. mays		
FEATURE:		
NAME/KEY: misc feature		
LOCATION: (1) (1198)		
OTHER INFORMATION: Ceres Seq. ID no. 12373616		
US-11-096-568A-19558		
Query Match	26.3%	Score 287.4; DB 9; Length 1198;
Best Local Similarity	62.4%	Pred. No. 3.3e-64;
Matches 450; Conservative 0; Mismatches 271; Indels 0; Gaps 0;		
QY	286	CAGCAGTTCGACGAAATTCGCTTTCGATGACGTCGACGACCGTCGACGCTTGGTG 345
DB	372	CGGCGCGCGCGCGCGACGCGTGTGCGAGCGAGCGCGTGTGCTGCTGACCGCGCG 431
QY	346	GAGCATAGCGCGTGTCTTAAGCTTCGAGATTCACCAAGGAAAGCTGATTCACAGA 405
DB	432	GCGCTACTCTCTTGTGCGCGCTTCGACGAGCTCACGAGCGCGGCTGATCGAAGA 491
QY	406	GTTTGACGAGAAAGCTTTCGATATCTTCGAGCTCTTTCGATTCGTCGTCGACCA 465
DB	492	GCTTGACGAGAAAGCTTTCGATATCTTCGAGCTCTTTCGATTCGTCGTCGACCA 551
QY	466	TCTTCAGGAGATTCGACGAGCTGATCTTTCGCTTTCGCTTTCGCTTTCGATGAGT 525
DB	552	TGTTGACGAAATTCGACGAGCTGATTCGCTTTCGCTTTCGCTTTCGCTTTCGATG 611
QY	526	TAAAGCTTGTATTAAGGAGCTATTCATTCGCCAATTCGATGCTATCAATCGTCA 585

```

Db      612 TGAAGCTTCTGATATATAGGACTCCGCTCTCACTAGATGAAGCTGTGGTAAATCAGTGA 671
Qy      586 CAAGAAGAGGAGACGAGAAAGTTCCTTAAAGTCCCTTTGTTCTAGCTTTAGCTCTTC 645
Db      672 CACGTGAAGGAAAAACGAGAGAAATTCCTGAGAGGCCACTATATATATCTCTGTGCTGC 731
Qy      646 TTTTCTGCGGTTTTCTTCTGAGAGAGTCTCCATCGATGATGATCTGTTAGCAATGA 705
Db      732 TGTTCAGCGTTTATGATCTTCTGCGGTAGTATCCCATGGGATGCTCTCTTGATGA 791
Qy      706 TGTGTGTGGCGATGGAATAGCTGATATATATGGAACGTAACTTTGGGTCACTAAGATAC 765
Db      792 TGAGCGGTGGGAGTGTGTTGCTGACATTTGTTGGAGAGGATATGAGCTCAGCAAGCTGC 851
Qy      766 CTTCACACCCAAAGAGTGTGGGACAGAACATCTCATGTTCATCTTCGCTTTCTCA 825
Db      852 CATTCAACTGGAGAGAGAGCTGGGCGGAGACATCTCATGTTCATTTCTGTTCCTGC 911
Qy      826 TCTCATCGCATTACTTACTATTACTCAAGCCTTGGGATCCTTCAACATGAACGGAGAA 885
Db      912 TGTTCGCGATATATGCTCTACTTCTCAAGCCTGGTTACATTTGATGATTTATCTGGAG 971
Qy      886 CGACCTTGACAGAGTACCAATGCTCTCAATGTGCGCACGGTACGTGAGTGCCTACCA 945
Db      972 AGGACTGTGTAAGCTGCGCTGTGTGCACTAGCAGACAGTAGTGAAGTGCCTCTG 1031
Qy      946 TCACCGATCAATTAAGACAAATATTTGGTTCCTCGGTACTATTTTACTGCTTAT 1005
Db      1032 TGACGAAGTTGTAGATGACAAATATCTGTTCTTTGCGCACATGCTGTAGCTTTTC 1091
Qy      1006 T 1006
Db      1092 T 1092

```

RESULT 2

US-11-096-568A-14317

; Sequence 14317, Application US/11096568A

; Publication No. US20060048240A1

; GENERAL INFORMATION:

; APPLICANT: Alexandrov, Nikolai et al.

; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides

; FILE REFERENCE: 2750-1592PUS2

; CURRENT APPLICATION NUMBER: US/11/096,568A

; NUMBER OF SEQ ID NOS: 34471

; SEQ ID NO 14317

; LENGTH: 481

; TYPE: DNA

; ORGANISM: Zea mays subsp. mays

; FEATURE:

; NAME/KEY: misc.feature

; LOCATION: (1)..(481)

; OTHER INFORMATION: Ceres Seq. ID no. 3626194

US-11-096-568A-14317

Query Match 15.5%; Score 169; DB 9; Length 481;

Best Local Similarity 63.8%; Pred. No. 9e-34;

Matches 256; Conservative 0; Mismatches 145; Indels 0; Gaps 0;

```

Qy      606 GAGTGTCTTAAAGTCTTTGTTCTACGTTCTAGCTCTTTCTCTGCGGTTTTCTTC 665
Db      26 GAATTCCTGAGAGGCTCACTATATATGCTCTGCTGCTCTTTAGAGCTTTTCTTC 85
Qy      666 TGGAGAGATCTCTATCGGTATGATCTGTTAGCAATGATGTGTGGCGAGTAA 725
Db      86 TGGCGTGAAGTCCCATCGGATCGCTCTGTCGATGATGAGCGGTGCGATGTTT 145
Qy      726 GCTGATATATAGGACGTAAAGTGTGGGTCACTAAGATACCTTACACCCAAAGAGT 785
Db      146 GCTGACATTTGTTGGAGAGGATATGCTCAGCAAGCTGCCATTCATCGGAAGAGC 205

```

RESULT 3

US-11-096-568A-1769

; Sequence 1769, Application US/11096568A

; Publication No. US20060048240A1

; GENERAL INFORMATION:

; APPLICANT: Alexandrov, Nikolai et al.

; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides

; FILE REFERENCE: 2750-1592PUS2

; CURRENT APPLICATION NUMBER: US/11/096,568A

; NUMBER OF SEQ ID NOS: 34471

; SEQ ID NO 1769

; LENGTH: 986

; TYPE: DNA

; ORGANISM: Zea mays subsp. mays

; FEATURE:

; NAME/KEY: misc.feature

; LOCATION: (1)..(986)

; OTHER INFORMATION: Ceres Seq. ID no. 15178973

US-11-096-568A-1769

Query Match 15.5%; Score 169; DB 9; Length 986;

Best Local Similarity 63.8%; Pred. No. 1.2e-33;

Matches 256; Conservative 0; Mismatches 145; Indels 0; Gaps 0;

```

Qy      606 GAGTGTCTTAAAGTCTTTGTTCTACGTTCTAGCTCTTTCTCTGCGGTTTTCTTC 665
Db      380 GAATTCCTGAGAGGCTCACTATATATGCTCTGCTGCTGTTCAAGGTTTTAGCTTC 439
Qy      666 TGGAGAGATCTCTATCGGTATGATCTGTTAGCAATGATGTGTGGCGAGTGA 725
Db      440 TGGCGTGAAGTCCCATCGGATCGCTCTGTCGATGATGAGCGGTGCGATGTTT 499
Qy      726 GCTGATATATAGGACGTAAAGTGTGGGTCACTAAGATACCTTACACCCAAAGAGT 785
Db      500 GCTGACATTTGAGAGAGGATATGCTCAGCAAGCTGCCATTCATTCGGAAGAGC 559
Qy      786 TGGCAGAGAACATCTCCATGTTCTATCTTCGCTTCTTCATCTCCATGCACTTACTTAC 845
Db      560 TGGCGCGGAGACATCTGATTTCACTTCTGTTTCCGCTGTCGCGATGATGCTC 619
Qy      846 TATTACTCAAGCTTGGTATCTTCACTGAATGGAAGAGACCTTGCAAGATGAC 905
Db      620 TACTTCTCAAGCTTGGTATCTTCACTTATGATGTTATCTGGGAAGAGGCACTGTGAGCTGCG 679
Qy      906 ATGCTCAATGTCGCGACGGTAGTGAAGTGCATCCATCAGCATCAATTAAGACGAC 965
Db      680 CTGTGCACTAGACGACAGTAGTGAAGTGCATCTGTGACCGAAGTTGATGAC 739
Qy      966 AATATTCGTTCTCTGCTACTATTTTACTGCTTATTT 1006
Db      740 AATATCTGTTCTTTGGCAGCATGCTGTAGCTTTCT 780

```


RESULT 4
 US-11-096-568A-14811
 ; Sequence 14811, Application US/11096568A
 ; Publication No. US20060048240A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Alexandrov, Nikolai et al.
 ; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
 ; TITLE OF INVENTION: Thedy
 ; FILE REFERENCE: 2750-1592PUS2
 ; CURRENT APPLICATION NUMBER: US/11/096,568A
 ; CURRENT FILING DATE: 2005-04-01
 ; NUMBER OF SEQ ID NOS: 34471
 ; SEQ ID NO 14811
 ; LENGTH: 1048
 ; TYPE: DNA
 ; ORGANISM: Zea mays subsp. mays
 ; NAME/KEY: misc_feature
 ; LOCATION: (1)..(1048)
 ; OTHER INFORMATION: Ceres Seq. ID no. 12336653
 US-11-096-568A-14811

Query Match 14.1%; Score 153.8; DB 9; Length 1048;
 Best Local Similarity 51.9%; Pred. No. 1e-29;
 Matches 372; Conservative 0; Mismatches 342; Indels 3; Gaps 1;

Qy 304 CGCTGTGATGAGCTGGAGCCCGTGGAGCTTGTGGAGCATAGCCGCTTGTCT 363
 Db 201 CCTGGCCCATGACCTGAGGCTGTCAGCTGTCACAGGTGGTGGCCCTGCTCTCA 260
 Qy 364 TAAGCTTCGAGAGCTCCACCAAGGAAAGCTATTCAACAGATTGGAGAAAGCTTG 423
 Db 261 AATTCTTGGAGAGCTGCCCAAGCGCGGCTTTCGAGCAAGAACTAGAGAAAGCTTG 320
 Qy 424 TGCAATATCTCTCAGGTCTGCTTTTGTGTACTTGGTGCGCAATCTTCAAGCGATCGA 483
 Db 321 TTCAATATAGAGCTTGGGCTGGATTAATGCTTTTGGCCCTTTTATAGCTGATAGT 380
 Qy 484 AGGCTGATATCTTGTGCTTTTGTTCGTTAGTGAATGCTTAAGCTTGTATTACG 543
 Db 381 ATACTCTTTCCTTGGCGCTGACACAGGGGTTAACTATTATAGAGTCTTCTACTGG 440
 Qy 544 GACTATCATATTTCCCAATTCAGATGTAATCAATCCGTCAAGAGAGAGAGAGAG 603
 Db 441 GGGTGGAGCTTATATAAAGAGAGAGCTATGTTAATCAATGAGCCGCTTGGAGATTACA 500
 Qy 604 AAGAGTGTCTTAAAGTCTTGTCTTCTAGCTTCTTCTTCTGCGGTTTCT 663
 Db 501 GGGAACTTTCAGAGGCGCACTGATTAATGCTGCTACTATTAATCTTTGCCAGTCTCTT 560
 Qy 664 TCTGGAGAGAGTCTCTATGCTATGATCTGCTTGAAGATGATGTGTGGCATGGA 723
 Db 561 TATGGAAGAGCTCAACAGTGTCTATAGCACTTATATCACTTATGTGCTGGGAGTGA 620
 Qy 724 TAGGTATATTAATGAGAGCTAATGTTGGGCACTAATAGATACCTTACCAACCAAGAA 783
 Db 621 TAGCCATGTTGTTGGAGAGAGCTTATGTAAGAAAGCTTCAATACCAATCCCAAAAT 680
 Qy 784 GTTGGGAGAGAGATCTCAATGTTTCACTTCTGCTTCTTCACTCATGCAATTACTT 843
 Db 681 CATATGCTGGAAGATAGCAATGCGGTGGCTGCTTCTTCTGCTGCTGATATGTC 740
 Qy 844 ACTATTAATCAAGCTTGGGTAAGCTTCAATGAACTGGGAAAGCACTTGGAGAGATAG 903
 Db 741 ATTACTTCAACACTTGGGTTTATGAGAAAGATGATGATGCTTT--AAGCTTCC 797
 Qy 904 CAATGCTCAATGAGTGGCCAGGTAAGTCAAGTCACTACCACTACCAATTAAGAG 963
 Db 798 TCGTGGCTCTGTAAGTCAAGTCACTGTTGAATGCAACCCCATAGACAGAAATTGATG 857
 Qy 964 ACAATATTTGGTCTCTGCTGCTAATTTTATAGTGTCTTATTTAAGTTTGGATATT 1020
 Db 858 ACAATTTGAGTGTCTTCTTATGATCATCTTCTAGTTGGAGCTCATCTTTTGAATTT 914

RESULT 5
 US-10-995-561-13236
 ; Sequence 13236, Application US/10995561
 ; Publication No. US20050272054A1
 ; GENERAL INFORMATION:
 ; APPLICANT: CARGILL, Michele et al.
 ; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
 ; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
 ; TITLE OF INVENTION: DETECTION AND USES THEREOF
 ; FILE REFERENCE: CL001559
 ; CURRENT APPLICATION NUMBER: US/10/995,561
 ; CURRENT FILING DATE: 2004-11-24
 ; NUMBER OF SEQ ID NOS: 85702
 ; SOFTWARE: PASCSEQ for Windows Version 4.0
 ; SEQ ID NO 13236
 ; LENGTH: 305312
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; NAME/KEY: misc_feature
 ; LOCATION: (1)..(305312)
 ; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-
 US-10-995-561-13236

Query Match 3.8%; Score 41; DB 8; Length 305312;
 Best Local Similarity 49.4%; Pred. No. 13;
 Matches 130; Conservative 1; Mismatches 131; Indels 1; Gaps 1;

Qy 6 AAGATAATATTAACAATAATATATTTCTTATCTTATGATGATGCAAGATTCTTCTTC 65
 Db 270457 AAGTTGTTACAAAGAGATCTTACATCATGATGATGACATTATTTTGTCTT 270516
 Qy 66 TTCTTCTTCTCTCTCTCTCAAACTCACTCCCTCGTCATGAGAGCAACCTTACT 125
 Db 270517 TCTCTTCTTCTCTCTCTCTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 270576
 Qy 126 CTATC-TCCATCAATATATAGTTGTGTGGGAAACACTTTGACAGTCAACG 184
 Db 270577 CTTCCTTCT 270636
 Qy 185 GTTCGTCTCTCTGAGCTTCTTGAATTTCTTCTCTCTTCTTCTTCTTCTTCTTCT 244
 Db 270637 TTCCCTGCTCTACCCCTCTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 270696
 Qy 245 CTCGTACTCACTTACGTCCTC 267
 Db 270697 GGTACTACTCACTTAAATGATC 270719

RESULT 6
 US-11-136-527-7472
 ; Sequence 7472, Application US/11136527
 ; Publication No. US2005028570A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Wyeth
 ; APPLICANT: Mounes, William M
 ; TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes
 ; FILE REFERENCE: 031896-041000 (AM101086)
 ; CURRENT APPLICATION NUMBER: US/11/136,527
 ; CURRENT FILING DATE: 2005-05-25
 ; PRIOR APPLICATION NUMBER: US 60/574,294
 ; PRIOR FILING DATE: 2005-05-26
 ; NUMBER OF SEQ ID NOS: 362830
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 7472
 ; LENGTH: 1400
 ; TYPE: DNA
 ; ORGANISM: Rattus norvegicus
 US-11-136-527-7472

Query Match 3.7%; Score 40.8; DB 12; Length 1400;

QY 255 CAGTTACGTGCTCGTCTCTGATCTTTCAGCAG 291
| | | | | | | | | | | | | | | | | | |
Db 319 TCTACACTGTTAGTCATTTTCTATCTCTCTGCTG 283

```

RESULT 10
US-09-925-065A-655076/C
; Sequence 655076, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT FILING DATE: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq For Windows Version 4.0
; SEQ ID NO 655076
; LENGTH: 571
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-655076

```

	Query Match	3.5%	Score 38.6	DB 6	Length 571
	Best Local Similarity	51.6%	Pred. No. 4.7		
	Matches 112	Conservative 1	Mismatches 100	Indels 4	Gaps 1
QY	79	TCTCTCTCCAAACTCAGTTCCTCCGTCGCATGACGACCACTTACCTCTATCTCCGATCA	138		
Db	499	TTCTTCCTTCGATCTTACAGATCTCTTACTCAATGCGCAAAATGCAAGAGCTTACAGATTC	440		
QY	139	ATCATCAAGTTATGTCGGGTTGGGGAACAATCTTTACAGATCAGCGGTTTCGTTCC-----T	194		
Db	439	ATTAATGGCGAGGCTACCTCTCGGTTTAAAGCATATAATTAACAAGTCATTTGGGCGAGGT	380		
QY	195	CTGGGCTTCTGATTTCTTCTCTCTGTTTCATTTGATTTGACCGGATGGGCTCTGCTACT	254		
Db	379	CATGGGTTGCTGACTCTAGGCGCTGGTTCTATTTTTCAGAACCTTCTGCTCTCTACT	320		
QY	255	CAGTTACTGCTGCTGCTGTTCTCTATCTCTTTCAGAG	291		
Db	319	TCTACAACTGTAACTATTTCTATATCTCTCTGCTG	283		

```

1 RESULT 11
2 US-11-143-401-45/c
3 ; Sequence 45, Application US/11143401
4 ; Publication No. US20060002953A1
5 ; GENERAL INFORMATION:
6 ; APPLICANT: Behr, Marcel
7 ; APPLICANT: Small, Peter
8 ; APPLICANT: Schoolnik, Gary
9 ; APPLICANT: Wilson, Michael A.
10 ; TITLE OF INVENTION: Molecular differences between Species of
11 ; TITLE OF INVENTION: the M. Tuberculosis Complex
12 ; FILE REFERENCE: STAN102CON
13 ; CURRENT APPLICATION NUMBER: US/11/143,401
14 ; CURRENT FILING DATE: 2005-06-01
15 ; PRIOR APPLICATION NUMBER: US/10/647,089
16 ; PRIOR FILING DATE: 2003-08-21
17 ; PRIOR APPLICATION NUMBER: US/09/894,844
18 ; PRIOR FILING DATE: 2001-06-27
19 ; PRIOR APPLICATION NUMBER: 09/318,191
20 ; PRIOR FILING DATE: 1999-05-25
21 ; PRIOR APPLICATION NUMBER: 60/097,936
22 ; PRIOR FILING DATE: 1998-08-25
23 ; NUMBER OF SEQ ID NOS: 137

```

```

1 SOFTWARE: FastSeq for Windows Version 4.0
2 SEQ ID NO 45
3 LENGTH: 395
4 TYPE: DNA
5 ORGANISM: Mycobacteria tuberculosis
6 FEATURE:
7 NAME/KEY: misc feature
8 LOCATION: (1)...(395)
9 OTHER INFORMATION: n = A,T,C or G
10 FEATURE:
11 NAME/KEY: misc feature
12 LOCATION: 27, 44, 104, 119, 180, 224, 237, 245, 254, 301, 327, 370,
13 LOCATION: 385, 393
14 OTHER INFORMATION: n = A,T,C or G
15 US-11-143-401-45

```

[illegible]

```

RESULT 12
US-09-925-065A-512623
: Sequence 512623, Application US/09925065A
: Publication No. US20040181048A1
: GENERAL INFORMATION:
: APPLICANT: Wang, David G.
: TITLE OR INVENTION: Identification and Mapping of Single
: TITLE OR INVENTION: Nucleotide Polymorphisms in the Human Genome
: FILE REFERENCE: 108827.135
: CURRENT APPLICATION NUMBER: US/09/925,065A
: CURRENT FILING DATE: 2001-08-08
: PRIOR APPLICATION NUMBER: US 60/243,096
: PRIOR FILING DATE: 2000-10-24
: PRIOR APPLICATION NUMBER: US 60/252,147
: PRIOR FILING DATE: 2000-11-20
: PRIOR APPLICATION NUMBER: US 60/250,092
: PRIOR FILING DATE: 2000-11-30
: PRIOR APPLICATION NUMBER: US 60/261,766
: PRIOR FILING DATE: 2001-01-16
: PRIOR APPLICATION NUMBER: US 60/289,846
: PRIOR FILING DATE: 2001-05-09
: NUMBER OF SEQ ID NOS: 957086
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 512623
: LENGTH: 605
: TYPE: DNA
: ORGANISM: Homo sapiens
US-09-925-065A-512623

```

Query Match 3.5%; Score 37.8; DB 6; Length 605;
Beet Local Similarity 52.8%; Pred. No. 7.7;
Matches 102; Conservative 1; Mismatches 88; Indels 2; Gaps 1;

Oy 56 TCTCTTCTCTCTCTCTCTCTCCTCCAAACTGATTGGTCGCATCGAGAC 115
db 315 TCCTCTCTCTCTCTTTCTCTCTCCCACTCCCTCTCTCTCTCTCTCCACACTCT 374

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: March 16, 2006, 19:49:20 ; Search time 24 Seconds
(without alignments)
1218.746 Million cell updates/sec

Title: US-10-634-548-2
Perfect score: 1536
Sequence: 1 MAATLPLSPINHQLCRFGNN.....DDNISVPLATILAAVLSFGY 304

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1536	100.0	304	2 T48445	hypothetical prote
2	195	12.7	233	2 S76857	hypothetical prote
3	191	12.4	190	2 G70433	hypothetical prote
4	184.5	12.0	211	2 D69413	conserved hypotet
5	167.5	10.9	246	2 A69013	conserved hypotet
6	164.5	10.7	179	2 AD2007	hypothetical prote
7	153.5	10.0	290	2 S58323	probable membrane
8	149	9.7	215	2 T45229	probable F420H2-de
9	144.5	9.4	519	2 J00124	membrane protein S
10	143.5	9.3	203	2 H75102	hypothetical prote
11	142.5	9.3	202	2 C71084	hypothetical prote
12	139.5	9.1	204	2 D71648	hypothetical prote
13	131.5	8.6	204	2 C97866	hypothetical prote
14	126	8.0	204	2 C97866	hypothetical prote
15	123	8.0	215	2 H90141	probable dolichol
16	119	7.7	257	2 C71265	conserved hypotet
17	114.5	7.5	191	2 H64441	hypothetical prote
18	113.5	7.4	207	2 D70151	conserved hypotet
19	112	7.3	456	2 H96902	membrane protein,
20	109	7.1	396	2 F69813	multidrug-efflux t
21	105.5	6.9	218	2 T40365	conserved hypotet
22	104.5	6.8	183	2 F69049	conserved hypotet
23	104.5	6.8	653	2 C82580	oligopeptide trans
24	104	6.8	598	2 T48980	hypothetical prote
25	103	6.7	545	2 S59143	NADH2 dehydrogenas
26	102.5	6.7	405	2 AC1193	hypothetical prote
27	102	6.6	654	2 T14202	NADH2 dehydrogenas
28	100.5	6.5	432	2 F63903	hypothetical prote
29	100.5	6.5	527	2 T03427	hypothetical prote

30	100.5	6.5	550	2 AF3244	conserved hypotet
31	100	6.5	521	2 S74569	NADH2 dehydrogenas
32	99.5	6.5	309	2 A82944	ribose/galactose A
33	99.5	6.5	373	2 AB1223	ethanolamine-utili
34	98	6.4	328	2 H84061	ferriochrome ABC tr
35	98	6.4	404	2 AF0034	probable integral
36	97.5	6.3	290	2 T06856	hypothetical prote
37	97.5	6.3	500	2 H95276	probable transmemb
38	97	6.3	430	2 P70016	purine permease ho
39	97	6.3	430	2 C97160	uracil permease ur
40	97	6.3	551	2 D71969	l-lactate permease
41	96.5	6.3	633	2 S82153	PTS system, fructo
42	95.5	6.2	288	2 G64102	phosphatidate cycl
43	95.5	6.2	373	2 AE1576	E. coli ethanolami
44	95.5	6.2	460	2 A39938	phosphotransferase
45	95.5	6.2	551	2 B64537	L-lactate permease

ALIGNMENTS

RESULT 1
T48445
hypothetical protein T32M21.90 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 09-Jul-2004
C:Accession: T48445
R:Bevan, M.; Terry, N.; Ardiss, W.; Buysbaert, C.; Daseville, R.; De Clerck, R.; De
ewes, H.W.; Rudd, S.; Lemcke, K.; Mayer, K.F.X.
submitted to the Protein Sequence Database, March 2000
A:Reference number: Z24487
A:Accession: T48445
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-304 <BEV>
A:Cross-references: UNIPROT:Q9LZ76; UNIPARC:UP100000A81C7; EMBL:AL162875
A:Experimental source: cultivar Columbia; BAC clone T32M21
C:Genetics:
A:Map position: 5
A:introns: 99/3; 122/2; 166/2; 205/1; 243/2
A:Note: T32M21.90

Query Match 100.0%; Score 1536; DB 2; Length 304;
Best Local Similarity 100.0%; Pred. No. 3.5e-119;
Matches 304; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAATLPLSPINHQLCRFGNNSLTTHRCSPGFLISSPCFGLTGMGSAQTQARRSLISS 60
DB 1 MAATLPLSPINHQLCRFGNNSLTTHRCSPGFLISSPCFGLTGMGSAQTQARRSLISS 60

QY 61 AATNSLHDVGAIVAVLGGAYALVSPESITKRNVIQOISRLQVHILSGLLFVLWPI 120
DB 61 AATNSLHDVGAIVAVLGGAYALVSPESITKRNVIQOISRLQVHILSGLLFVLWPI 120

QY 121 FSGSTEARYPFAFPLVNLGLRLVINGLSISPSNLSIVTREGRAEELKGPLFYVALT 180
DB 121 FSGSTEARYPFAFPLVNLGLRLVINGLSISPSNLSIVTREGRAEELKGPLFYVALT 180

QY 181 FSAVFFWRRESPIGMSISLMMCGGDIADIMGKRGSTKIPNPRKSNAGSISMEIFGFFI 240
DB 181 FSAVFFWRRESPIGMSISLMMCGGDIADIMGKRGSTKIPNPRKSNAGSISMEIFGFFI 240

QY 241 SIALLYTSSISGLYHNMWETTLQVAVNSVATVYESLPITDQDDNISVPLATILAAVL 300
DB 241 SIALLYTSSISGLYHNMWETTLQVAVNSVATVYESLPITDQDDNISVPLATILAAVL 300

QY 301 SFGY 304
DB 301 SFGY 304

RESULT 2
S76857

hypothetical protein - *Synechocystis* sp. (strain PCC 6803)

C:Species: *Synechocystis* sp.

A:Variety: PCC 6803

C:Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 09-Jul-2004

C:Accession: S76857

R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;

O, K.; Okumura, S.; Shimizu, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yaeuda

DNA Res. 3, 109-136, 1996

A:Title: Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis*

A:Reference number: S74322; MUID:97061201; PMID:8905231

A:Accession: S76857

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-233 <KAN>

A:Cross-references: UNIPROT:P74653; UNIPARC:UPI000000C09EC; EMBL:D90917; GB:AB001339; NID

A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

Query Match

Best Local Similarity 29.1%; Score 195; DB 2; Length 233;

Matches 69; Conservative 49; Mismatches 83; Indels 36; Gaps 12;

QY

71 VGATVAVGAVYALVLSFSLTKRNVIOQSRLVHLISGLFVLA---PIFGSGTE 126

DB 17 VGLATVYIG---AVLTALRLNRLSLSPAETRKIVHIGAGVVLIAWMLSPGVGAI- 72

QY

127 ARYFAFPLVNGRLVNGLSISPNMLKSVTEGRAEELKGLPYVAL-LFSAVF 185

DB 73 AGVFPA-----GIVLSRLPLPS---LBSVGHNS-----YETLVYALISGLVGGF 117

QY

186 FWRSEPI-GMISLAWCGGDIADIMGRKFGSTKIP-YNPKSMAGSISMPFIFGISIA 243

DB 118 FSLGPIPAIGLIVMAGDDGIALVVGGRGHRQVGFGRKSMGTLTIVLASPLVTVV 177

QY

244 LLYYSSISGLYHNMWETTLQRYAMVSMVATVESIPITDQD-DNISVPLATILAAY 299

DB 178 FLSY--TRGF-----TVIVLVAGTVAIASAGLESFRMGIDMLTVPLGSALIAW 225

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB 112 FSLGVYFGRKRLFYNPKSLBETLAFPTASF---IGLLLF-----TDFCEA 156

QY

266 AMVSMATVVSLEPIPTDQDNLISVPLATIAAYL 300

DB

157 FVLSICAVLESPL--KDDNRYIP---VLASPL 186

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY 75 VAVLGAVALVLSPESTLTKRVIOQSLSRKLVHLSG-LFVLAMPFSGSTEARVPAF 133
 DB 43 VAVL-----LVSRFRFQDR-----PNLSRKFVHLMVGNILFL--PLFESRLVTTFLLAA- 90
 QY 134 VPLVNGRLVINGLSISPNMILKSVTRREGRAEELKGPLFYVALLPFAVFWRESPIG 193
 DB 91 APFL-----LITFL-LSPYSPFLRVKRASSYGHL--GLVYYSISMTILATLFFEARWIT 142
 QY 194 MISLMMCGDDGDIADIMGRKFGSTKIP-YNPRKSWAGSISMFIFGPFISIALLYYSSLG 252
 DB 143 GIGIAAASYGDFASLTGERFRGRTTFSVLGDCKSLKESLGFMIFILLVWLPLVLTYYSG-Q 201
 QY 253 YLHMNMTTLQORVAMVAVATVVSLEPTDQDINISPLATILAAVSPG 303
 DB 202 F-----TPYLLIGVALVSTVLEAL--TPRGDNLNTRACGAV-AAAYILLG 242

RESULT 6

AD2007
 hypothetical protein all1610 [imported] - Nostoc sp. (strain PCC 7120)
 C/Species: Nostoc sp. PCC 7120
 A/Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
 C/Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
 C/Accession: AD2007
 R/Kaneho, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Matanabe, A.; Iriyuchi, Nakazaki, N.; Shimo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S
 DNA Ref: 8, 205-213, 2001
 A/Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena
 A/Reference number: AB1807; MUID:21595285; PMID:11759840
 A/Accession: AD2007
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-179 <KUR>
 A/Cross-references: UNIPROT:Q8YWK3; UNIPARC:UPI000000CE149; GB:BA000019; PIDN:BA877976.1;
 A/Experimental source: strain PCC 7120
 C/Genetics:
 A/Gene: all1610

Query Match 10.7%; Score 164.5; DB 2; Length 179;
 Best Local Similarity 27.5%; Pred. No. 2e-06;
 Matches 53; Conservative 40; Mismatches 71; Indels 29; Gaps 8;
 QY 113 LFVLAMPF-----FSGSTEARVPAFVPLVNGRLVINGLSISPNMILKSVTRREGRAEEL 169
 DB 1 MILLAMWLDIPASVIGASVVASIVTLLSYLFPILPG-----INSVERQS----- 45
 QY 170 KGPLFYVALLPFAVFW--RESPIGMISLMMCGDDIADIMGRKFGSTKIP-YNPRKS 226
 DB 46 LGTFYVAVSVGLVAMFRHIQPOYAAIGMMVAMKGDIALVQGRFGKRYKALLGAKKS 105
 QY 227 WAGSISMFIFGPFISIALLYYSSLGVLHMNMTTLQORVAMVAVATVVSLEPTDQD 286
 DB 106 WEGSLVTWMAISYVCSIL-----LGVLGVMQVTLVSLA-VAVFATSLKAFSLGV--D 157
 QY 287 NISVPLATILAAV 299
 DB 158 NLTVPLGSAIAIAF 170

RESULT 7

558323
 probable membrane protein YOR311c - yeast (Saccharomyces cerevisiae)
 N/Alternate names: hypothetical protein O611
 C/Species: Saccharomyces cerevisiae
 C/Date: 13-Jan-1996 #sequence_revision 01-Mar-1996 #text_change 09-Jul-2004
 C/Accession: 558323; S67217; S71991
 R/Pearson, B.M.; Hernandez, Y.; Wolf, S.S.; Kalogeropoulos, A.; Schweizer, M.
 submitted to the EMBL Data Library, August 1995
 A/Reference number: 558318
 A/Accession: 558323
 A/Molecule type: DNA
 A/Residues: 1-290 <PEA>

A/Cross-references: UNIPROT:Q12382; UNIPARC:UPI000006B14; EMBL:X90565; NID:G940836; PID
 R/Pearson, B.M.; Hernandez, Y.; Kalogeropoulos, A.; Schweizer, M.
 submitted to the Protein Sequence Database, July 1996
 A/Reference number: S67213
 A/Accession: S67217
 A/Molecule type: DNA
 A/Residues: 1-290 <PEM>
 A/Cross-references: UNIPARC:UPI000006B14; EMBL:Z75219; NID:G1420683; PID:e252144; PID:G
 A/Experimental source: strain S288C
 R/Pearson, B.M.; Hernandez, Y.; Payne, J.; Wolf, S.S.; Kalogeropoulos, A.; Schweizer, M.
 Yeast 12, 1021-1031, 1996
 A/Title: Sequencing of a 35.71 kb DNA segment on the right arm of yeast chromosome XV re
 A/Reference number: S71986; MUID:97051589; PMID:8896266
 A/Accession: S71991
 A/Status: nucleic acid sequence not shown; translation not shown
 A/Molecule type: DNA
 A/Residues: 1-290 <PEP>
 A/Cross-references: UNIPARC:UPI000006B14; EMBL:X90565; NID:G940836; PIDN:CA62166.1; PI
 A/Note: the nucleotide sequence was submitted to the EMBL Data Library, August 1995
 C/Genetics:
 A/Gene: SCD:HSD1
 A/Cross-references: SCD:S0005838
 A/Map position: 15R
 A/Keywords: transmembrane protein
 F/101-117/Domain: transmembrane #status predicted <TM1>
 F/146-162/Domain: transmembrane #status predicted <TM2>
 F/200-216/Domain: transmembrane #status predicted <TM3>
 F/243-259/Domain: transmembrane #status predicted <TM4>

Query Match 10.0%; Score 153.5; DB 2; Length 290;
 Best Local Similarity 26.3%; Pred. No. 2.9e-05;
 Matches 61; Conservative 42; Mismatches 78; Indels 51; Gaps 10;

QY 80 GAVLVLSFESLTKRVNIOQSLSRKLVHLSGLFVLAMPFSGSTEARVPAFVPLVNG 139
 DB 84 GFILVLYTOGINKVKNVLMPLIYAFILFLIDIL-RINWPF----- 124
 QY 140 LRLVINGLSISPNMILKSVTRREGRAEEL--KGPLFYVALLPFAVFWRESPIGMISL 197
 DB 125 -----NMLYKRTVAGLMRKKEIHTYNAVLTLLGLTFSPNFSKD--VTLISL 170
 QY 198 AMCGGDIADIMGRKFG--STKIPYNPRKSWAGSISMFIFGPFISIALLYYSSLGVLH 255
 DB 171 FILMSDTAAATIGRKXGHLTPKAR--KSLAGSIAFTVG-VITCMVFQGVPAVSY 227
 QY 256 MN-----W--ETTLQORVAMVAVATVVSLEPTDQ--DDNISVPLATIL 296
 DB 228 VNKRGEIOWSPETSRILSNMLSLGVAALSEGIDLFNMDNFTIPLVSL 279

RESULT 8

T45229
 probable F420H2-dehydrogenase [imported] - Methanobolus tindarius
 N/Alternate names: ffd protein
 C/Species: Methanobolus tindarius
 C/Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 09-Jul-2004
 R/Westendberg, D.U.; Braune, A.; Ruppert, C.; Mueller, V.; Herzberg, C.; Gottschalk, G.;
 submitted to the EMBL Data Library, September 1998
 A/Description: The F420H2-dehydrogenase from Methanobolus tindarius: Cloning of the ffd
 A/Reference number: Z22947
 A/Accession: T45229
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: DNA
 A/Residues: 1-215 <WES>
 A/Cross-references: UNIPROT:Q9UXN7; UNIPARC:UPI00000624B8; EMBL:AJ011519; PIDN:CA856642.
 A/Experimental source: DSM 2278
 C/Genetics:
 A/Gene: ffd

Query Match 9.7%; Score 149; DB 2; Length 215;
 Best Local Similarity 24.3%; Pred. No. 4.7e-05;
 Matches 53; Conservative 45; Mismatches 66; Indels 54; Gaps 10;

Oy 103 RCLVHLISGLIFPLAMPFISGSGTEARYPAAVPLVNGRLVINGISISPNMLTIS----- 158
 Db 17 RCLVHVISGLIAYI---PLVYFTEISLIVFL-----LEPWF-PTIVSLARAKTOP 67
 Oy 159 ----VIREGRAEE---LLKGPLFYVLLALFSFAVFWRESPIGMI--SLAMMCGDGIADI 209
 Db 68 VCDMIRKWRKRNKKNYLPLKATLLINTGILSYFLF-----PNIVCAALATLGGGIAIV 123
 Oy 210 MGRKFGSTKYIPNPRKSMAGSISMFIFGF-----FISIALLYYSSIGLYLM 256
 Db 124 AGERFEFKHNPYSERKRYEGTMAGIIAFASTVIVSPVALFGSVGALFIESVIG---- 179
 Oy 257 MNETTLQRVAMSGMVAIVESLPTDQLDNDINSPLAT 294
 Db 180 ---RDLRTGSSINSFINLIK-----NDNLVLEPIAS 206

 RESULT 9
 J00124
 membrane protein SEC59 - yeast (Saccharomyces cerevisiae)
 N/Aternate names: protein YM6270.17c; protein YMR013c
 C/Species: Saccharomyces cerevisiae
 C/Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 09-Jul-2004
 C/Accession: J00124; S53044
 R.Bernstein, M. Kepes, F. J. Schekman, R.
 Mol. Cell. Biol. 9, 1191-1199, 1989
 A/Title: SEC59 encodes a membrane protein required for core glycosylation in Saccharomyces
 A/Accession number: J00124; MUID:89261723; PMID:2657387
 A/Molecule type: DNA
 A/Residues: 1-519 <BER>
 A/Cross-references: UNIPROT:P20048; UNIPARC:UPI00001355EA; EMBL:M25779; NID:g172567; PIR
 R.Devlin, K. J. Churcher, C.M. March 1995
 submitted to the EMBL Data Library, March 1995
 A/Reference number: S53028
 A/Accession: S53044
 A/Molecule type: DNA
 A/Residues: 1-519 <DEV>
 A/Cross-references: UNIPARC:UPI00001355EA; EMBL:Z48613; NID:g728645; PID:g728662; MIPS:K
 A/Experimental source: strain AB972
 C/Comment: This protein acts either directly or indirectly in the biosynthesis of dolich
 C/Geneticus:
 A/Gene: SGD:SEC59
 A/Cross-references: SGD:S0004615; MIPS:YMR013c
 A/Map position: 13R
 C/Keywords: glycoprotein; transmembrane protein
 F/87-103/Domain: transmembrane #status predicted <TM1>
 F/158-174/Domain: transmembrane #status predicted <TM2>
 F/181-197/Domain: transmembrane #status predicted <TM3>
 F/229-245/Domain: transmembrane #status predicted <TM4>
 F/260-276/Domain: transmembrane #status predicted <TM5>
 F/301-317/Domain: transmembrane #status predicted <TM6>
 F/397-413/Domain: transmembrane #status predicted <TM7>
 F/452-468/Domain: transmembrane #status predicted <TM8>
 F/503-519/Domain: transmembrane #status predicted <TM9>
 F/326/Binding site: carbohydrate (Aen) (covalent) #status predicted

 Query Match 9.4%; Score 144.5; DB 2; Length 519;
 Best Local Similarity 25.3%; Pred. No. 0.00031;
 Matches 57; Conservative 51; Mismatches 74; Indels 43; Gaps 12;

 Oy 84 LVLSFES-LTKRANYIOGSLRKLVIHLISGLIFPLAMPFISGSGTEARYPAAVPLVNGRL 142
 Db 309 LIISIPILILEKDSLSTNSRKLHMF---IIFLLIIPSTQMS-----NFKVI 353
 Oy 143 VING---LSISPNMLIKSVTREGRAEEL-----KQPLFY-VLALLF-SAVFF 186
 Db 354 ALSGTIPVPLSIEYIRFQNLPLPGALIEQLRFRADDRDHSGLPILISYLILFGISTPL 413
 Oy 187 WRSSPIGMISLAMMCGDGIADIMGRKFGSTKYIPNPRKSMAGSISMFIFGPISIALY 246
 Db 414 MNNSPMLIGIGI---GDSLSIIGKRYGRIRNK- GTOKTLGGLAFIVTSFIVCLVLR 469

[illegible]

Db 1 MSKLKELRKALHMTGLSVPLIYALGKSAITPVLIIL-----FLALBPRLGE 52
 Qy 139 GLRLVINGLSISPNML-----IKSVTRREGAEELKGPLFYVALALLPSAVFWRSP 192
 Db 53 GLRLVKEKGLIPESVTEKIEREIDGIARE-REKGIQAHYFTIALLLVLYLFPREVAI 111
 Qy 193 GMISLMMCGDGDINDIMGRKFGSTKYIPYPRKSMAGSISMFIQGFISIALLYYSSLSIG 252
 Db 112 GSISVATL--GDMAAIIIGKSGRHR--FKNGKSVESGLAFITILLITPLV----- 160
 Qy 253 YLHNMETTLQRVANVNAVATVE--SLPTDQLDNDISVPLATILAAVLS 301
 Db 161 -----GLKMAALIASLVGMIVVEFYGLP-----PDDNFSNOALAIITLYLA 199

RESULT 12

hypothetical protein RP860 - Rickettsia prowazekii
 N:Alternate names: hypothetical protein dnag 5'-region
 C:Species: Rickettsia prowazekii
 C>Date: 21-Nov-1998 #sequence_revision 21-Nov-1998 #text_change 09-Jul-2004
 C:Accession: D11648; A46754
 R:Anderson, S.G.E.; Zomorodipour, A.; Andersson, J.O.; Slicheritz-Ponten, T.; Alsmark, U.
 Nature 396, 133-140, 1998
 A:Title: The genome sequence of Rickettsia prowazekii and the origin of mitochondria.
 A:Reference number: A71630; MUID:99039499; PMID:9823893
 A:Accession: D11648
 A:Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-204 <AND>
 A:Cross-references: UNIPROT:P41078; UNIPARC:UPI0000139D6E; GB:AJ235273; GB:AJ235269; NID
 A:Experimental source: strain Madrid E
 R:Mark, G.L.; Wood, D.O.
 Gene 123, 121-125, 1993
 A:Title: Characterization of the gene coding for the Rickettsia prowazekii DNA primase
 A:Reference number: J0534; MUID:93138420; PMID:8422995
 C:Accession: A46754
 A:Status: not compared with conceptual translation
 A:Molecule type: DNA
 A:Residues: 177-204 <MAR>
 A:Cross-references: UNIPARC:UPI000017AB2C
 C:Genetics:
 A:Gene: RP860

Query Match 9.1%; Score 139.5; DB 2; Length 204;
 Best Local Similarity 26.9%; Pred. No. 0.00027;
 Matches 59; Conservative 41; Mismatches 74; Indels 45; Gaps 11;
 Qy 103 RKLVIHLSGLFVLAAMPFSGSTEARYPAPV--LVNGRLVINGLSI-----SPNSM 154
 Db 12 RKIFH-LSAIFPM-----FYLFVPRIALMLFITSTITLVDIRHNNAK 57
 Qy 155 LIKSVTR-----EGRAELKGPLFYVALALLPSAVFWRSPIGMISLMMCGDG 205
 Db 58 IRKVTTFPSKIRLKENNGTFALSGISFMMLGFFLSILPPKN--LVICSMILLIISDC 115
 Qy 206 IADIMGRKFGSTKYIPYPRKSMAGSISMFIQGFISIALLYYSSLSGLYHNMETTLQ 265
 Db 116 LAALVGIKIGS--LSNGKSIAGSTFPFVSALFISI-LVYFY--LGY-----NTSPVLI 164
 Qy 266 AMVSMVATVVESSLPTDQLDNDISVPLATILAAVLSFGY 304
 Db 165 IISCIQATVAEFYSKIRINDNISIPLSYCLSTTI-PPY 202

RESULT 13

C97866
 hypothetical protein RC1331 [imported] - Rickettsia conorii (strain Malish 7)
 C:Species: Rickettsia conorii
 C>Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 09-Jul-2004
 C:Accession: C97866
 R:Ogata, H.; Audic, S.; Reneato-Audiffren, P.; Fournier, P.E.; Barbe, V.; Samson, D.; Rd
 Science 293, 2093-2098, 2001

A:Title: Mechanisms of Evolution in Rickettsia conorii and Rickettsia prowazekii.
 A:Reference number: A97700; MUID:21442074; PMID:11557893
 A:Accession: C97866
 A:Status: Preliminary
 A:Molecule type: DNA
 A:Residues: 1-204 <KUR>
 A:Cross-references: UNIPROT:Q92FZ6; UNIPARC:UPI00000CC08B; GB:AE006914; PIDN:AA03869.1;
 C:Genetics:
 A:Gene: RC1331

Query Match 8.6%; Score 131.5; DB 2; Length 204;
 Best Local Similarity 26.0%; Pred. No. 0.0012;
 Matches 57; Conservative 43; Mismatches 72; Indels 47; Gaps 11;

Qy 103 RKLVIHLSGLFVLAAMPFSGSTEARYPAPV-----LVNGRLVINGLSISPN 153
 Db 12 RKIFH-LSAIFPLY-----LPIPTAMTLFIITATLYLD-VSRHNA 56
 Qy 154 MLISVTR-----EGRAELKGPLFYVALALLPSAVFWRSPIGMISLMMCGD 204
 Db 57 TISFVTRFPSKYRLERNNGSPALSGVSEFMWIGFFLTALPPKN--LVICSMILLIISD 114
 Qy 205 GIADIMGRKFGSTKYIPYPRKSMAGSISMFIQGFISIALLYYSSLSGLYHNMETTLQ 264
 Db 115 CLALVGIKIGS--LSNGKSIAGSTIFLASAIFISI-LVYFY--LGY-----NTSPVLI 163
 Qy 265 AMVSMVATVVESSLPTDQLDNDISVPLATILAAVLSF 302
 Db 164 IISCIQATVAEFYSKIRINDNISIPLSYCLSTALISY 202

RESULT 14

T41511
 Probable dolichol kinase - fission yeast (Schizosaccharomyces pombe)
 C:Species: Schizosaccharomyces pombe
 C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
 C:Accession: T41511
 R:Murphy, L.; Harris, D.; Wood, V.; Rajandream, M.A.; Barrell, B.G.
 submitted to the EMBL Data Library, March 1999
 A:Reference number: 221999
 A:Accession: T41511
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-465 <MR>
 A:Cross-references: UNIPROT:Q9Y7T6; UNIPARC:UPI0000069554; EMBL:AL049522; PIDN:CAE40014.
 A:Experimental source: strain 972h-; cosmid c63
 C:Genetics:
 A:Gene: SPDB:SPCC63.10C
 A:Map position: 3
 A:Introns: 77/2; 106/3; 118/2

Query Match 8.2%; Score 126; DB 2; Length 465;
 Best Local Similarity 19.9%; Pred. No. 0.0092;
 Matches 79; Conservative 51; Mismatches 112; Indels 154; Gaps 19;

Qy 17 FGNNSLTTR-----FCSGFLISSPCF--GLTNGSGATQPARRSLISAVAT 64
 Db 94 FGNSVLYHMYTMEALYCFRWGTFPIEGILCTGLTGLCPAT----- 138
 Qy 65 NSLHDVGAIVAVLGAVALVLSPESLTKGNVIGQSLSRK-----LVHLSGLLF 114
 Db 139 --PSYEVSPVSVLSG--VLISLPTILANCLIKLAALHLSALPTTCLIIYFSSALLV 194
 Qy 115 VL--AV-----PIFS-----GSTE 126
 Db 195 FLVSRVVAQGLQAPVWLFNQIFSHRNSLTRIKIIMWIIICGCFPIILSRNNPL 254
 Qy 127 ARYFAA-----FYPL-----VNGRLVING 146
 Db 255 GKYPTEDEVINFRKRYHALVFLFVPCCLDPHFLHSFGVLFIFLVEGRI--- 310
 Qy 147 LSIISPNMLIKSVTRREGRAELKGPLF--YVALPSAVFWM---RSGPIGMSI--A 198

Db 311 LRLKPGKMIHEFLMEYTDNRDHKPLIISHIYLLIGCAIPIMLSNALKGPVASELLVG 370
 QY 199 MMC--GGDGLDINGRKRGSTIKIPNPKKSWAGSISMIFGPFISIALIYYSSLGTYLM 256
 Db 371 VLICGCGSMASIIIGKRGKRIS-KTKKSIEG---VFASISVFLVHLTQA--FHV 422
 QY 257 NMETTLQRYAMVAMVAVESLPITDQDDNISVPL 292
 Db 423 CPSVTFWKTLFMSLCTALLEGVSTE--NDMLILPM 455

RESULT 15

H90141

conserved hypothetical protein [imported] - Sulfolobus solfataricus

C:Species: Sulfolobus solfataricus

C:Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 09-Jul-2004

C:Accession: H90141

R:She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Aweez, M.J.; Chan-

Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, R

arrett, R.A.; Ragan, M.A.; Sensen, C.W.; Van der Oost, J.

submitted to GenBank, April 2001

A:Description: Sulfolobus solfataricus complete genome.

A:Reference number: A9139

A:Accession: H90141

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-215 <KUR>

A:Cross-references: UNIPROT:Q981C1, UNIPARC:UPI0000064170, GB:AE006641, NID:g13813155; F

C:Gene: S800024

C:Superfamily: Aeropyrum pernix hypothetical protein APE2054

Query Match 8.0%; Score 123; DB 2; Length 215;

Best Local Similarity 25.4%; Pred. No. 0.0066;

Matches 54; Conservative 41; Mismatches 66; Indels 52; Gaps 12;

QY 101 LSRKLVHL-SGLFLVLAIPFGSTEARYPAAFPVPLVNGRLVINGLSISPSNMLIKSV 159
 Db 37 VTRKAIHMLGGIVAVLSPFVN-----SPLVPITTSYIMAYLITI---RVLKRM 85
 QY 160 TREGRAEELKQPLFYVIALLPSAVF-----FWRESP--IGMISLAMCGGDGIA 207
 Db 86 SWFMOKENL--GEVVF--AFSFGVILVMFILDPMYWTSTDLIYAVVPLIFMSFGDGIT 141
 QY 208 DIMGRKSGSTIKIPNPKKSWAGSISMIFGPFISIALIYYSSLGTYLMNMETTLQRYAM 267
 Db 142 GII-RNF-----VYRR-----VKGLGSMGMLIFCATVGFVNLIPGLT----- 180
 QY 268 VSMVATVESLPITDQDDNISVPLATIIAAYL 300
 Db 181 AGIATVABLDP---YIDDNISVPFLSFLFLYL 210

Search completed: March 16, 2006, 19:51:29

Job time : 27 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: March 16, 2006, 19:47:55 ; Search time 110 Seconds
(without alignment)

1949.824 Million cell updates/sec

Title: US-10-634-548-2
Perfect score: 1536
Sequence: 1 MAATLPLSPINHQCRCGNM.....DDNISVPLATLAAVLSFGY 304

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues
Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt 05.80:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1536	100.0	304	09L276_ARATH	09L276 arabidopsis
2	1531	99.7	304	084WCO_ARATH	084WCO arabidopsis
3	1523	99.2	304	08LF61_ARATH	08LF61 arabidopsis
4	791.5	51.5	314	07XR51_ORYSA	07XR51 oryza sativ
5	565.5	36.8	246	05N9J9_ORYSA	05N9J9 oryza sativ
6	531.5	34.6	307	067ZM7_ARATH	067ZM7 arabidopsis
7	530.5	34.5	307	067ZM1_ARATH	067ZM1 arabidopsis
8	526.5	34.3	307	08LAW9_ARATH	08LAW9 arabidopsis
9	455.5	29.7	517	09LV00_ARATH	09LV00 arabidopsis
10	309.5	20.1	1144	086HKE_DICDI	086HKE dictyosteli
11	285	16.6	223	055A10_DICDI	055A10 dictyosteli
12	257	16.7	223	050PDE_ENTHI	050PDE entamoeba h
13	245.5	16.0	230	050VQ1_ENTHI	050VQ1 entamoeba h
14	234	15.2	237	08KB08_CHLITE	08KB08 chlorobium
15	230	15.0	230	050LW4_ENTHI	050LW4 entamoeba h
16	195	12.7	233	P74653_SYNY3	P74653 synecocyst
17	191	12.4	190	067497_AQUAE	067497 aquifex aeo
18	184.5	12.0	211	028960_ARCFU	028960 archaeoglob
19	180	11.7	312	06KZ85_PICTO	06KZ85 picophyllus
20	173	11.3	317	075317_ASHGO	075317 ashbya goss
21	167.5	10.9	246	027171_METTH	027171 methanobact
22	164.5	10.7	179	08YWK3_ANASP	08YWK3 anabaena sp
23	163	10.6	315	0978U8_THEVO	0978U8 thermoplasm
24	161	10.4	217	07V9G9_PROMA	07V9G9 proma
25	159	10.4	283	06CKR7_KLUUA	06CKR7 kluyveromyc
26	158	10.3	204	073J58_TREDE	073J58 treponema d
27	157	10.2	218	08DHCO_SYNEB	08DHCO synecococc
28	156.5	10.2	217	05N4S4_SYNP6	05N4S4 synecococc
29	156	10.2	217	07V3U3_PROMM	07V3U3 prochloroco
30	153.5	10.0	290	012382_YEAST	012382 saccharomyc
31	149	9.7	215	09UXN7_METTI	09UXN7 methanobabu

32	146	9.5	216	2	07UD5_SYNPX	07UD5 synecococc
33	145.5	9.5	193	2	08TMQ4_METTA	08TMQ4 methanopyru
34	144.5	9.4	519	1	SECS5_YEAST	P20048 saccharomyc
35	143.5	9.3	203	2	09UZD4_PPRAB	09UZD4 pyrococcus
36	143	9.3	223	2	08Q0A6_METWA	08Q0A6 methanosarc
37	142.5	9.3	202	2	058680_PYRHO	058680 pyrococcus
38	142.5	9.3	227	2	07NGM2_GLOVI	07NGM2 gloeobacter
39	140	9.1	281	2	06FO20_CANCA	06FO20 candida gla
40	139.5	9.1	204	1	Y860_RICPR	P41078 rickettsia
41	139.5	9.1	384	2	05YIM5_CRYNE	05YIM5 cryptococcu
42	136	8.9	225	2	08TPA4_METAC	08TPA4 methanosarc
43	135	8.8	204	1	06HVQ3_RICTY	06HVQ3 rickettsia
44	134	8.7	282	2	09HLC3_THEAC	09HLC3 thermoplasm
45	132.5	8.6	216	2	04JIC4_SULFO	04JIC4 sulfolobus

RESULT 1	09L276_ARATH	PRELIMINARY;	PRT;	304 AA.
ID	09L276;			
AC	09L276;			
DT	01-OCT-2000 (TREMBLrel. 15, Created)			
DT	01-OCT-2000 (TREMBLrel. 15, Last sequence update)			
DT	10-MAY-2005 (TREMBLrel. 30, Last annotation update)			
DE	Hypothetical protein 132M21_90 (Ac5904490).			
GN	Name=132M21_90;			
OS	Arabidopsis thaliana (Mouse-ear cress).			
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;			
OC	eucosids II; Brassicales; Brassicaceae; Arabidopsids.			
OX	NCBI_TaxID=3702;			
RN	[1]			
RP	NUCLEOTIDE SEQUENCE.			
RA	Bevan M., Terryn N., Ardiles W., Buysbaert C., Dasseville R.,			
RA	De Clerck R., De Keyser A., Neyt P., Rouze P., Van Den Daele H.,			
RA	Villarroel R., Gielens J., Van Montagu M., Bancroft I., Mewes H.W.,			
RA	Rudd S., Lemcke K., Mayer K.F.X.;			
RL	Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.			
RN	[2]			
RP	NUCLEOTIDE SEQUENCE.			
RA	EU Arabidopsis sequencing project;			
RL	Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.			
RN	[3]			
RP	NUCLEOTIDE SEQUENCE.			
RA	Kim C.J., Chen H., Cheuk R., Shim P., Ecker J.R.;			
RT	"Arabidopsis ORF clones."			
RL	Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AL162875; CAB85555.1; -; genomic_DNA.			
DR	EMBL; BT021123; AAX2258.1; -; mRNA.			
DR	PIR; T48445; T48445.			
DR	GO; GO:0016020; C:membrane; IEA.			
DR	GO; GO:0004605; F:phosphatidate cytidyllyltransferase activity; IEA.			
DR	GO; GO:0008654; P:phospholipid biosynthesis; IEA.			
DR	InterPro; IPR000374; PC_trans.			
DR	Pfam; PF01148; CTP_transf_1; 1.			
KW	Hypothetical protein.			
SQ	SEQUENCE 304 AA; 33089 MW; 770CA569C9F50A50 CRC64;			

Query Match	100.0%;	Score 1536;	DB 2;	Length 304;
Best Local Similarity	100.0%;	Pred. No. 7.9e-112;		
Matches 304;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

QY	1	MAATLPLSPINHQCRCGNNSLTTHRCSPGFLISSPCFGLTGWGSAATQARARSLISS	60
DB	1	MAATLPLSPINHQCRCGNNSLTTHRCSPGFLISSPCFGLTGWGSAATQARARSLISS	60
QY	61	AAATNSLHDVGAATVAATLGGAAVAVLSFESITKRNVIQOSISRLCLVHLSGLVFLAMP	120
DB	61	AAATNSLHDVGAATVAATLGGAAVAVLSFESITKRNVIQOSISRLCLVHLSGLVFLAMP	120
QY	121	FGSGTEARVFAAPVLVNGRLVINGLSISPSNMLISVTRGRABELLKGPLFYVIALLL	180

Db 121 FSGSTEARYFAFVPLVNGLRVINGLSISPSNMLIKSVTREGRAEELKGPLFVLAALL 180
Qy 181 FSAVFFMRBESPIGMISLAMCGGDIADIMGRKFGSTIKIPNPRKSWAGSISMFTFGFFI 240
Db 181 FSAVFFMRBESPIGMISLAMCGGDIADIMGRKFGSTIKIPNPRKSWAGSISMFTFGFFI 240
Qy 241 SIALLYSSISGLYLMHNETTLQRYAVMSVAVTVESLPITDQDDNISVPLATILAAVL 300
Db 241 SIALLYSSISGLYLMHNETTLQRYAVMSVAVTVESLPITDQDDNISVPLATILAAVL 300
Qy 301 SFGY 304
Db 301 SFGY 304

RESULT 2

Q84WC0 ARATH PRELIMINARY; PRT; 304 AA.
AC Q84WC0;
DT 01-JUN-2003 (Tremblrel. 24, Last sequence update)
DT 01-JUN-2003 (Tremblrel. 24, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE Hypothetical protein At5g04490.
GN Name=At5g04490;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eustoids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP NUCLEOTIDE SEQUENCE.

RA Yamada K., Chan M.M., Chang C.H., Dale J.M., Hsuan V.W., Lee J.M.,
RA Onodera C.S., Quach H.L., Tang C., Toriumi M., Wong C., Wu H.C.,
RA Yu G., Yuan S., Carninci P., Chen H., Cheuk R., Hayashizaki Y.,
RA Ishida J., Jones T., Kamita A., Kawai J., Kim C.J., Narusaka M.,
RA Nguyen M., Palm C.J., Sakurai T., Satou M., Seki M., Shim P.,
RA Southwick A., Trupp M.G., Wu T., Shinozaki K., Davis R.W., Ecker J.R.,
RA Theologis A.;
RA Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
RL EMBL; BT004006; AA042044.1; -, mRNA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0004605; F:phosphatidate cytidyltransferase activity; IEA.
DR GO; GO:0008654; P:phospholipid biosynthesis; IEA.
DR InterPro: IPR000374; PC:trans.
DR Pfam: PF01148; CTP_transf_1; 1.
KW Hypothetical protein.
SQ SEQUENCE 304 AA; 33059 MW; 361DBFC3635FA7FB CRC64;

Query Match 99.7%; Score 1531; DB 2; Length 304;

Best Local Similarity 99.7%; Pred. No. 1.9e-111;
Matches 303; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MAATPLSPINHQLCRFGNNSLTTHRFCSPPGLISSPCFGLTGWGSATQLRARRSLISS 60
Db 1 MAATPLSPINHQLCRFGNNSLTTHRFCSPPGLISSPCFGLTGWGSATQLRARRSLISS 60
Qy 61 AVATNSLHDVGATVAVVGAVALVLSFESLTGRNVIOQSLSRKLVHLSGLFVLAMP 120
Db 61 AVATNSLHDVGATVAVVGAVALVLSFESLTGRNVIOQSLSRKLVHLSGLFVLAMP 120
Qy 121 FSGSTEARYFAFVPLVNGLRVINGLSISPSNMLIKSVTREGRAEELKGPLFVLAALL 180
Db 121 FSGSTEARYFAFVPLVNGLRVINGLSISPSNMLIKSVTREGRAEELKGPLFVLAALL 180
Qy 181 FSAVFFMRBESPIGMISLAMCGGDIADIMGRKFGSTIKIPNPRKSWAGSISMFTFGFFI 240
Db 181 FSAVFFMRBESPIGMISLAMCGGDIADIMGRKFGSTIKIPNPRKSWAGSISMFTFGFFI 240
Qy 241 SIALLYSSISGLYLMHNETTLQRYAVMSVAVTVESLPITDQDDNISVPLATILAAVL 300
Db 241 SIALLYSSISGLYLMHNETTLQRYAVMSVAVTVESLPITDQDDNISVPLATILAAVL 300

Qy 301 SFGY 304
Db 301 SFGY 304

RESULT 3

Q8LF61 ARATH PRELIMINARY; PRT; 304 AA.
AC Q8LF61;
DT 01-OCT-2002 (Tremblrel. 22, Created)
DT 01-OCT-2002 (Tremblrel. 22, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE Hypothetical protein.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eustoids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP NUCLEOTIDE SEQUENCE.

RA MEDLINE=22088475; PubMed=12093376;
RA Haas B.J., Volkovskiy N., Town C.D., Troukhan M., Alexandrov N.,
RA Feldmann K.A., Flavell R.B., White O., Salzberg S.L.;
RA "Full-length messenger RNA sequences greatly improve genome
RA annotation.";
RL Genome Biol. 3:RESEARCH0029-RESEARCH0029 (2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Brover V., Troukhan M., Alexandrov N., Lu Y.-P., Flavell R.,
RA Feldmann K.;
RA Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
RL EMBL; AY085036; AA061593.1; -, mRNA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0004605; F:phosphatidate cytidyltransferase activity; IEA.
DR GO; GO:0008654; P:phospholipid biosynthesis; IEA.
DR InterPro: IPR000374; PC:trans.
DR Pfam: PF01148; CTP_transf_1; 1.
KW Hypothetical protein.
SQ SEQUENCE 304 AA; 33069 MW; 937B18D809F3A626 CRC64;

Query Match 99.2%; Score 1523; DB 2; Length 304;

Best Local Similarity 98.7%; Pred. No. 8.2e-111;
Matches 300; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MAATPLSPINHQLCRFGNNSLTTHRFCSPPGLISSPCFGLTGWGSATQLRARRSLISS 60
Db 1 MAATPLSPINHQLCRFGNNSLTTHRFCSPPGLISSPCFGLTGWGSATQLRARRSLISS 60
Qy 61 AVATNSLHDVGATVAVVGAVALVLSFESLTGRNVIOQSLSRKLVHLSGLFVLAMP 120
Db 61 AVATNSLHDVGATVAVVGAVALVLSFESLTGRNVIOQSLSRKLVHLSGLFVLAMP 120
Qy 121 FSGSTEARYFAFVPLVNGLRVINGLSISPSNMLIKSVTREGRAEELKGPLFVLAALL 180
Db 121 FSGSTEARYFAFVPLVNGLRVINGLSISPSNMLIKSVTREGRAEELKGPLFVLAALL 180
Qy 181 FSAVFFMRBESPIGMISLAMCGGDIADIMGRKFGSTIKIPNPRKSWAGSISMFTFGFFI 240
Db 181 FSAVFFMRBESPIGMISLAMCGGDIADIMGRKFGSTIKIPNPRKSWAGSISMFTFGFFI 240
Qy 241 SIALLYSSISGLYLMHNETTLQRYAVMSVAVTVESLPITDQDDNISVPLATILAAVL 300
Db 241 SIALLYSSISGLYLMHNETTLQRYAVMSVAVTVESLPITDQDDNISVPLATILAAVL 300

Qy 301 SFGY 304
Db 301 SFGY 304

RESULT 4

Q7KR51 ORYSA PRELIMINARY; PRT; 314 AA.
ID Q7KR51 ORYSA
AC Q7KR51;

DT 01-OCT-2003 (TReMBLrel. 25, Created)
 DT 01-OCT-2003 (TReMBLrel. 25, Last sequence update)
 DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
 DE OSUNBa0043A12.34 protein.
 GN Name=OSUNBa0043A12.34;
 OS Oryza sativa (japonica cultivar-group).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Erihartoideae; Oryzaceae; Oryza.
 NC NCB1_TaxID=399477;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RX PubMed:12447439; DOI=10.1038/nature01183;
 RA Feng Q., Zhang Y., Hao P., Wang S., Pu G., Huang Y., Li Y., Zhu J.,
 Li Y., Hu X., Jia P., Zhang Y., Zhao Q., Ying K., Yu S., Tang Y.,
 Weng Q., Zhang L., Lu Y., Mu Y., Lu Y., Zhang L.S., Yu Z., Fan D.,
 Lu X., Lu T., Li C., Wu Y., Sun T., Lei H., Li T., Hu H., Guan J.,
 Wu M., Zhang R., Zhou B., Chen Z., Chen L., Jin Z., Wang R., Yin H.,
 Cai Z., Ren S., Lv G., Gu W., Zhu G., Tu Y., Jia Y., Zhang Y.,
 Chen J., Kang H., Chen X., Shao C., Sun Y., Hu Q., Zhang X., Zhang W.,
 Wang L., Ding C., Sheng H., Gu J., Chen S., Ni L., Zhu F., Chen W.,
 Lan L., Lai Y., Cheng Z., Gu W., Jiang J., Li J., Hong G., Xue Y.,
 Han B.;
 RT "Sequence and analysis of rice chromosome 4.";
 RL Nature 420:316-320(2002).
 DR EMBL; AL666619; CAB02829.1; -; Genomic_DNA.
 DR Gramene; Q7XR51; -;
 DR GO; GO:0016020; C:membrane; IEA.
 DR GO; GO:0004605; P:phosphatidate cytidyltransferase activity; IEA.
 DR GO; GO:0008654; P:phospholipid biosynthesis; IEA.
 DR InterPro; IPR000374; PC_trans.
 DR Pfam; PF01148; CTP_transf_1; 1.
 SQ SEQUENCE 314 AA; 33391 MW; 97B941B926576CF8 CRC64;
 Query Match 51.5%; Score 791.5; DB 2; Length 314;
 Best Local Similarity 51.1%; Pred. No. 1.2e-53;
 Matches 159; Conservative 60; Mismatches 81; Indels 11; Gaps 4;
 QY 2 AATPLPSINQLCR---FGNNSL---TTRFGSPGLISSPCTIGT-GKGSATQLRA 53
 DB 3 AAAPFVDVRRPFCSSSVAASSSLISRSKRLSPAAAAASMKRRRLVLTGVGAAA--A 59
 QY 54 RRLSSAVATNSLHDVCAATVAVGAYALVSPESLTGKNNVIOQSRLVHLISGL 113
 DB 60 PAVAAALASATPAALRDCAATLLITAGAYSLVRAFDGLARLLLEQNSKRIYHLISGL 119
 QY 114 FYVALPFSGSTEARFPAFVPLVNGRLVINGLISPSNMLIKSVTEGRAEELIKGPL 173
 DB 120 FMSWMPLEFSNSTEARFPAFVPLVNGRLVINGLISPSNMLIKSVTEGRAEELIKGPL 179
 QY 174 FYVALPFSAVFPMRESPIGIMISLAMCGGDIADIMGRKGTSTYKIPNPKSMAAGSISM 233
 DB 180 YVIVLVLVSVVFWKQSPIGIVSLSMGSGDGFADIVGRYGSALCPFNENKSMWGISISM 239
 QY 234 FIFGFFSIALLYYSSIGYLMNMWETTLORVAVSMVATVSESPIPDOLDNISVPLA 293
 DB 240 FIFGFFSIALMLYFSSIGYLTTCVCDLALGKALVALVAVVECLPVDDVDNISVPLA 299
 QY 294 TILAAYLSFGY 304
 DB 300 TMLAAVLLFGY 310
 RESULT 5
 Q5N9J9_ORYZA PRELIMINARY; PRT; 246 AA.
 AC Q5N9J9;
 DT 01-FEB-2005 (TReMBLrel. 29, Created)
 DT 01-FEB-2005 (TReMBLrel. 29, Last sequence update)
 DT 01-FEB-2005 (TReMBLrel. 29, Last annotation update)
 DE Hypothetical protein P0446G04.41.
 GN Name=P0446G04.41;
 OS Oryza sativa (japonica cultivar-group).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Erihartoideae; Oryzaceae; Oryza.
 NC NCB1_TaxID=399477;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA Sasaki T., Matsuno T., Yamamoto K., Sakata K., Baba T., Katayose Y.,
 Wu J., Nishimura Y., Cheng Z., Nagamura Y., Antonio B.A., Kanamori H.,
 RA Hosokawa S., Maekawa M., Arikawa K., Chiden Y., Hayashi M.,
 RA Okamoto M., Ando T., Aoki H., Arita K., Hamada M., Harada C.,
 RA Hijishita S., Honda M., Ichikawa Y., Idonuma A., Iijima M., Ikeda M.,
 RA Ikono M., Itoh S., Itoh T., Itoh Y., Itoh Y., Iwabuchi A., Kamiya K.,
 RA Karasawa W., Kasegiri S., Kikuta A., Kobayashi N., Kono I.,
 RA Machida K., Maehara T., Mizuno H., Mizubayashi T., Mukai Y.,
 RA Nagasaki N., Nakashima M., Nakama Y., Nakamichi Y., Nakamura M.,
 RA Namiki N., Negishi M., Ohta I., Ono N., Saji S., Sakai K., Shibata M.,
 RA Shimokawa T., Shimura A., Song J., Takasaki Y., Teraawa K., Teuji K.,
 RA Waki K., Yamagata H., Yamane H., Yoshiki S., Yoshitake R., Yukawa K.,
 RA Zhong H., Iwama H., Endo T., Ito H., Hahn J.H., Kim H.I., Bun M.Y.,
 RA Yano M., Jiang J., Gojobori T.;
 RT "The genome sequence and structure of rice chromosome 1.";
 RL Nature 420:312-316(2002).
 DR EMBL; AP003252; BAB1857.1; -; Genomic_DNA.
 DR GO; GO:0016020; C:membrane; IEA.
 DR GO; GO:0004605; P:phosphatidate cytidyltransferase activity; IEA.
 DR GO; GO:0008654; P:phospholipid biosynthesis; IEA.
 DR InterPro; IPR000374; PC_trans.
 DR Pfam; PF01148; CTP_transf_1; 1.
 DR Hypothetical protein.
 SQ SEQUENCE 246 AA; 26457 MW; F830AB7300BE24E8 CRC64;
 Query Match 36.8%; Score 565.5; DB 2; Length 246;
 Best Local Similarity 44.5%; Pred. No. 4.1e-36;
 Matches 106; Conservative 47; Mismatches 84; Indels 1; Gaps 1;
 QY 65 NSLHDVCAATVAVGAYALVSPESLTGKNNVIOQSRLVHLISGLFYLAAMPFSGS 124
 DB 10 SGLAHDGSAAVTVAGVALALRPFEBLAKGVFPQKRNKLVHTIGVFLPLPFPFSG 69
 QY 125 TEARYFAFVPLVNGRLVINGLISPSNMLIKSVTEGRAEELIKGPLFYVALLES 184
 DB 70 SYAPFLAAVAPGINIMRLGLGVNMKNEMAVKMSRSGDPRRLKGPLYVATITTFATS 129
 QY 185 FPMRESPIGIMISLAMCGGDIADIMGRKGTSTYKIPNPKSMAAGSISMIFGFFSIAL 244
 DB 130 IFWRTSPALALINLCAAGDIADIVGRIGQELCPNPNKSYGSIAMALAGFMAISGY 189
 QY 245 LYYSSIGYLMNMWETTLORVAVSMVATVSESPIPDOLDNISVPLATILAAYLSF 302
 DB 190 MHYFQSPFIBESNLSAF-GLVAVSVTALVAVSHPISTHLDNLTVPULTSLVGSIVF 246
 RESULT 6
 O67ZW7_ARATH PRELIMINARY; PRT; 307 AA.
 AC O67ZW7;
 DT 25-OCT-2004 (TReMBLrel. 28, Created)
 DT 25-OCT-2004 (TReMBLrel. 28, Last sequence update)
 DT 25-OCT-2004 (TReMBLrel. 28, Last annotation update)
 DE Hypothetical protein At5g58560.
 GN Name=At5g58560;
 OS Arabidopsis thaliana (mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 NC NCB1_TaxID=3702;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA Totsuki Y., Seki M., Ishida J., Nakajima M., Enju A., Kamiya A.,
 RA Narusaka M., Shin-i T., Nakagawa M., Sakamoto N., Oishi K., Kohara Y.,
 RA Kobayashi M., Toyoda A., Sakaki Y., Sakurai T., Iida K., Akiyama K.,
 RA Saito M., Toyoda T., Konagaya A., Carninci P., Kawai J.,
 RA Hayashizaki Y., Shinzeki K.;

RT "Large-scale analysis of RIKEN Arabidopsis full-length (RAFL) cDNAs."
 RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AK176090; BAD43853.1; -; mRNA.
 DR GO: GO:0016020; C:membrane; IEA.
 DR GO: GO:0004605; F:phosphatidate cytidyltransferase activity; IEA.
 DR GO: GO:0008654; P:phospholipid biosynthesis; IEA.
 DR InterPro: IPR000374; PC:trans.
 DR Pfam: PF01148; CTP transf_1; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 307 AA; 33204 MW; 598BB44BD4450C1F CRC64;

Query Match 34.6%; Score 531.5; DB 2; Length 307;
 Best Local Similarity 38.1%; Pred. No. 2.4e-33;
 Matches 120; Conservative 66; Mismatches 108; Indels 21; Gaps 7;

QY 1 MAATLPSPINHQCRFGNNSLTTHRCSP---GFLISSPCFGL---TGMSATQLRAR 54
 DB 1 MATSTTKLSVLCSSFISSPLVD---SPSLAFPSPIRFLVRIATSRSSSRPPT 56
 QY 55 RSLISSAVAT---NSLHDV---GATVAVLGAYALVLSFSLTKRNVIQOISLRKVH 107
 DB 57 KIRSSLAAMVPENSVLSDVCAFGVTSIV---AFSCIGFWEIGKRGIPQKILRKLVH 113
 QY 108 ILSGLFLVLAAMPISGSTEARYPAFPVLVNGLRVLVNGLSISNSMLIKSVTEGRAE 167
 DB 114 INIGLVFMLCWPFLPSSGIQGLFASLVPGNLIVRMILLGLGVYHDEGTIKMSRRGRRE 173
 QY 168 LKGPFLVLAALFSAVFWRFPRESPIGMISLMMCGGDIADIMGRKFGSTKIPNPRKS 227
 DB 174 LKGPFLVLAALFSAVFWRFPRESPIGMISLMMCGGDIADIMGRKFGSTKIPNPRKS 233
 QY 228 AGSISMEIFGFEFISIALLYYSSLGYLMMWETTLQRYAMVSWATVESLPITDQDND 287
 DB 224 AGSISMAATAGFLASAVMYVYFASFGYIEDSGMIL-RFLVISIASALVESLPISDIDN 292
 QY 288 ISVPLATITLAAYLSF 302
 DB 293 LTISLTSALAGFLF 307

RESULT 7
 Q672A1_ARATH PRELIMINARY; PRT; 307 AA.
 AC Q672A1;
 DT 25-OCT-2004 (TREMBlrel. 28, Created)
 DT 25-OCT-2004 (TREMBlrel. 28, Last sequence update)
 DT 25-OCT-2004 (TREMBlrel. 28, Last annotation update)
 DE Hypothetical protein At5g358560.
 GN Name=At5g358560;
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsie.
 OC NCBI_TaxId=3702;
 RX NUCLEOTIDE SEQUENCE.
 RA Tsuchida Y., Seki M., Ishida J., Nakajima M., Enju A., Kamiya A.,
 RA Narusaka M., Shin-i T., Nakagawa M., Sakamoto N., Oishi K., Kohara Y.,
 RA Kobayashi M., Toyoda A., Sakaki Y., Sakurai T., Iida K., Akiyama K.,
 RA Satou M., Toyoda T., Konagaya A., Carninci P., Kawai J.,
 RA Hayashizaki Y., Shinozaki K.;
 RT "Large-scale analysis of RIKEN Arabidopsis full-length (RAFL) cDNAs."
 RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AK176217; BAD43980.1; -; mRNA.
 DR GO: GO:0016020; C:membrane; IEA.
 DR GO: GO:0004605; F:phosphatidate cytidyltransferase activity; IEA.
 DR GO: GO:0008654; P:phospholipid biosynthesis; IEA.
 DR InterPro: IPR000374; PC:trans.
 DR Pfam: PF01148; CTP transf_1; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 307 AA; 33218 MW; F98BB457D44D1079 CRC64;

Query Match 34.5%; Score 530.5; DB 2; Length 307;

Best Local Similarity 37.8%; Pred. No. 2.8e-33;
 Matches 119; Conservative 67; Mismatches 108; Indels 21; Gaps 7;

QY 1 MAATLPSPINHQCRFGNNSLTTHRCSP---GFLISSPCFGL---TGMSATQLRAR 54
 DB 1 MATSTTKLSVLCSSFISSPLVD---SPSLAFPSPIRFLVRIATSRSSSRPPT 56
 QY 55 RSLISSAVAT---NSLHDV---GATVAVLGAYALVLSFSLTKRNVIQOISLRKVH 107
 DB 57 KIRSSLAAMVPENSVLSDVCAFGVTSIV---AFSCIGFWEIGKRGIPQKILRKLVH 113
 QY 108 ILSGLFLVLAAMPISGSTEARYPAFPVLVNGLRVLVNGLSISNSMLIKSVTEGRAE 167
 DB 114 INIGLVFMLCWPFLPSSGIQGLFASLVPGNLIVRMILLGLGVYHDEGTIKMSRRGRRE 173
 QY 168 LKGPFLVLAALFSAVFWRFPRESPIGMISLMMCGGDIADIMGRKFGSTKIPNPRKS 227
 DB 174 LKGPFLVLAALFSAVFWRFPRESPIGMISLMMCGGDIADIMGRKFGSTKIPNPRKS 233
 QY 228 AGSISMEIFGFEFISIALLYYSSLGYLMMWETTLQRYAMVSWATVESLPITDQDND 287
 DB 224 AGSISMAATAGFLASAVMYVYFASFGYIEDSGMIL-RFLVISIASALVESLPISDIDN 292
 QY 288 ISVPLATITLAAYLSF 302
 DB 293 LTISLTSALAGFLF 307

RESULT 8
 Q8LAW9_ARATH PRELIMINARY; PRT; 307 AA.
 AC Q8LAW9;
 DT 01-OCT-2002 (TREMBlrel. 22, Created)
 DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
 DE Hypothetical protein.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsie.
 OC NCBI_TaxId=3702;
 RX NUCLEOTIDE SEQUENCE.
 RA Haas B.J., Volfovsky N., Town C.D., Troupkhan M., Alexandrov N.,
 RA Feldmann K.A., Flavell R.B., White O., Salberg S.L.;
 RT "Full-length messenger RNA sequences greatly improve genome
 annotation."
 GN Genome Biol. 3:RESEARCH0029-RESEARCH0029 (2002).
 GN [2]
 RN NUCLEOTIDE SEQUENCE.
 RA Brover V., Troupkhan M., Alexandrov N., Lu Y.-P., Flavell R.,
 RA Feldmann K.;
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AY087555; AA065097.1; -; mRNA.
 DR GO: GO:0016020; C:membrane; IEA.
 DR GO: GO:0004605; F:phosphatidate cytidyltransferase activity; IEA.
 DR GO: GO:0008654; P:phospholipid biosynthesis; IEA.
 DR InterPro: IPR000374; PC:trans.
 DR Pfam: PF01148; CTP transf_1; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 307 AA; 33204 MW; F98BB251D44B1679 CRC64;

Query Match 34.3%; Score 526.5; DB 2; Length 307;
 Best Local Similarity 37.5%; Pred. No. 5.8e-33;
 Matches 118; Conservative 67; Mismatches 109; Indels 21; Gaps 7;

QY 1 MAATLPSPINHQCRFGNNSLTTHRCSP---GFLISSPCFGL---TGMSATQLRAR 54
 DB 1 MATSTTKLSVLCSSFISSPLVD---SPSLAFPSPIRFLVRIATSRSSSRPPT 56
 QY 55 RSLISSAVAT---NSLHDV---GATVAVLGAYALVLSFSLTKRNVIQOISLRKVH 107

Db 57 KIRKSLAAMVPENSVLSDVCAFGVTSV---AFSCIFGFWGELGKRGIPDQKLRKLVH 113
 Qy 108 ILGLFLVLAAMPFSGSTEARFAFVPLVNLGRLVINGLSRPSMLIKSTREGRABE 167
 Db 114 INIGLVFVLCVPLFSSGIGQALFASLVPLVNLGLGVYHDEGTIKSMRGRDRE 173
 Qy 168 LKGLFLVLAALFSAVFWRSPICMISLMMCGGDDIADIMGRKFGSTKIPYPRKSM 227
 Db 174 LKGLFLVLAALFSAVFWRSPICMISLMMCGGDDIADIMGRKFGSTKIPYPRKSM 227
 Qy 228 AGSISMPFGEFFISIALLYSSIGLYLHMNETTLQRYAMVSVATVESPITDQDND 287
 Db 234 AGSISMAATGAFASVGYMYEFASFGYIEDSGMIL-RELVISIASALVESIPISDIDND 292
 Qy 288 ISVPLATLTAAYLSP 302
 Db 293 LTISLTSLAGPLFR 307

RESULT 9

Q9LV00 ARATH PRELIMINARY; PRT; 517 AA.
 AC Q9LV00;
 DT 01-OCT-2000 (Tremblrel. 15, Created)
 DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
 DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
 DE Emb|CAB8555.1.
 OS Arabidopsis thaliana (Mouse-ear cress).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 CC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 CX NCBI_TaxID=3702;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA MEDLINE=20181125; PubMed=10718197;
 RA Sato S., Nakamura Y., Kaneko T., Katoh T., Asamizu E., Kotani H.,
 RA Tabata S.;
 RT "Structural analysis of Arabidopsis thaliana chromosome 5. X. Sequence
 RT features of the regions of 3,076,755 bp covered by sixty P1 and TAC
 RT clones";
 RL DNA Res. 7:31-63(2000).
 RL EMBL; AB020755; BAA97326.1; -; Genomic DNA.
 SQ SEQUENCE 517 AA; 56646 MW; 85D803CE7688BE8A CRC64;

Query Match 29.7%; Score 455.5; DB 2; Length 517;
 Best Local Similarity 37.1%; Pred. No. 3.5e-27;
 Matches 106; Conservative 59; Mismatches 100; Indels 21; Gaps 7;

Qy 1 MAATLPLSPINHLQCRFGNNSLTTRFCSP---GFLISSPCIGL---TGMSATQLAR 54
 Db 1 MATSTTTKLSVLCSSFISSPLVD---SPSLAFSPFIPRLTVRIATSPRSSRPAT 56
 Qy 55 RSLISSAVAT---NSLHADV---GATVAVLGAYALVLSFESLTKRNVIOQSLRKLH 107
 Db 57 KIRKSLAAMVPENSVLSDVCAFGVTSV---AFSCIFGFWGELGKRGIPDQKLRKLVH 113
 Qy 108 ILGLFLVLAAMPFSGSTEARFAFVPLVNLGRLVINGLSRPSMLIKSTREGRABE 167
 Db 114 INIGLVFVLCVPLFSSGIGQALFASLVPLVNLGLGVYHDEGTIKSMRGRDRE 173
 Qy 168 LKGLFLVLAALFSAVFWRSPICMISLMMCGGDDIADIMGRKFGSTKIPYPRKSM 227
 Db 174 LKGLFLVLAALFSAVFWRSPICMISLMMCGGDDIADIMGRKFGSTKIPYPRKSM 227
 Qy 228 AGSISMPFGEFFISIALLYSSIGLYLHMNETTLQ-RVAMVSVATVESPITDQDND 287
 Db 234 AGSISMAATGAFASVGYMYEFASFGYIEDSGMIL-RELVISIASALVESIPISDIDND 292

RESULT 10
 Q86HK6 D1CD1 PRELIMINARY; PRT; 1144 AA.
 AC Q86HK6;

DT 01-JUN-2003 (Tremblrel. 24, Created)
 DT 01-JUN-2003 (Tremblrel. 24, Last sequence update)
 DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
 DE Similar to Oryza sativa (Japonica cultivar-group). P0446G04.23 protein
 DE (Fragment).
 OS Dictyostelium discoideum (Slime mold).
 CC Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.
 CX NCBI_TaxID=44689;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA STRAIN=AX4;
 RX MEDLINE=22092622; PubMed=12097910; DOI=10.1038/nature00847;
 RA Gloeckner G., Eichinger L., Szafrenski K., Pachebat J.A.,
 RA Barkley A.T., Dear P.H., Lehmann R., Baumgart C., Parra G.,
 RA Abtill J.F., Guigo R., Kumpf K., Tunggal B., Cox E., Quail M.A.,
 RA Platzner M., Rosenthal A., Noegel A.A.;
 RT "Sequence and analysis of chromosome 2 of Dictyostelium discoideum";
 RL Nature 418:79-85(2002).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RA STRAIN=AX4;
 RA Baumgart C.;
 RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AC116989; AAO52579.2; -; Genomic DNA.
 FT NON TER 1 1
 SQ SEQUENCE 1144 AA; 128163 MW; 748CFD3A53433730 CRC64;

Query Match 20.1%; Score 309.5; DB 2; Length 1144;
 Best Local Similarity 29.4%; Pred. No. 2.1e-15;
 Matches 90; Conservative 56; Mismatches 107; Indels 53; Gaps 11;

Qy 20 NSLTHPCSPGLISSPCFGLTGMGSAHQRLARRSLISAVALNLSLHVGATVAVLG 79
 Db 871 NSLVGNSPSSIKFQTSN-----NLDAIENNKTTISNSSITN-----GCQVSIKS 914
 Qy 80 GAYALVLSFESLTKR-----NVIQSLSRGLVITLSGLFLVLAAMPF-SGST 125
 Db 915 GHKSMALFNSCMERKSPVYTYIFSAVGVCIGL---LHIGGIIVLVNRRFPQCNW 971
 Qy 126 EARYFAFVPLVNLGRLVINGLSRPSMLIKSTREGRABEILKGLFLVLAALFSAVF 185
 Db 972 YSKIVGLVPLIISFQYALIGLIINDQKTVESMSRSGSPRELLGFLSYGIITSLTMI 1031
 Qy 186 FWRSPICMISLMMCGGDDIADIMGRKFGSTKIPYPRKSMASISMF---IGFFISI 242
 Db 1032 FW-RSPISILITIGVLCGDFPALFGLKYGKRIPIYRREKTLISLAFICSGTFTLL 1090
 Qy 243 AL-----LYYSSIGLYLHMNETTLQRYAMVSVATVESPITDQDNDISVPLATIAA 298
 Db 1091 TLQDRLLYPSIVLAPSLFW-----VCLISTLISLPLRW--DNITISICSVITL 1139
 Qy 299 YLSFGY 304
 Db 1140 NL-MGY 1144

RESULT 11

Q55A10 D1CD1 PRELIMINARY; PRT; 223 AA.
 AC Q55A10;
 DT 13-SEP-2005 (Tremblrel. 31, Created)
 DT 13-SEP-2005 (Tremblrel. 31, Last sequence update)
 DT 13-SEP-2005 (Tremblrel. 31, Last annotation update)
 DE Hypothetical protein.
 GN ORFNames=DD80216976;
 OS Dictyostelium discoideum (Slime mold).
 CC Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.
 CX NCBI_TaxID=44689;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA STRAIN=AX4;
 RA Eichinger L., Pachebat J.A., Gloeckner G., Rajandream M.-A.,
 RA Sugang R., Berriman M., Song J., Olsen R., Szafrenski K., Xu Q.,

RA Tunggal B., Kummerfeld S., Madera M., Konfortov B.A., Rivero F.,
 RA Bankier A.T., Lehmann R., Hamlin N., Davies R., Gaudet P., Fey P.,
 RA Plicher K., Chen G., Saunders D., Sodergren E., Davis P.,
 RA Kethornou A., Nie X., Hall N., Anjard C., Hemphill L., Bason N.,
 RA Farbrother P., Desany B., Just E., Morio T., Rost R., Churcher C.,
 RA Cooper J., Haydock S., van Driessche N., Cronin A., Goodhead I.,
 RA Murny D., Moutier T., Pain A., Lu M., Harper D., Lindsay R.,
 RA Hauser H., James K., Quiles M., Mohan M.B., Saito T., Buchrieser C.,
 RA Wardropet A., Felder M., Thangavelu M., Johnson D., Knights A.,
 RA Louisedge H., Mungall K., Oliver K., Price C., Quail M.A.,
 RA Urushihara H., Hernandez J., Rabinowitz E., Steffen D., Sanders M.,
 RA Ma J., Kohara Y., Sharp S., Simmons M., Spiegler S., Tivey A.,
 RA Sugano S., White B., Walker D., Woodward J., Winckler T., Tanaka Y.,
 RA Shualey G., Schieher M., Weinstein G., Rosenthal A., Cox E.C.,
 RA Chisholm R.L., Glibbe R., Loomis W.F., Platzer M., Kay R.R.,
 RA Williams J., Dear P.H., Noegel A.A., Barrell B., Kapa A.,
 RT "The genome of the social amoeba Dictyostelium discoideum.";
 RL Nature 0:0-0(2005).
 CC -!- CAUTION: The sequence shown here is derived from an
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
 CC preliminary data.
 CC DR EMBL; AAF0100020; EAL71342.1; -; Genomic_DNA.
 CC KW Hypothetical protein.
 SQ SEQUENCE 223 AA; 25192 MW; 83AB36930DF15C26 CRC64;

Query Match 18.6%; Score 285; DB 2; Length 223;
 Best Local Similarity 33.2%; Pred. No. 3e-14; Indels 22; Gaps 8;
 Matches 76; Conservative 44; Mismatches 87;

84 LVLSESLTKRVNIQOSTSRKLVHLSGLFVLAAMPFSGSTEARPAFVPLVNGRL 142
 ID LTLSTICLMTWMLNFCQFLKKHKVTS-STGIIVLVWRIFQFWYRSIVGLVLIISQY 67
 AC OS0506_1
 DT 13-SEP-2005 (TREMBLrel. 31, Last sequence update)
 DT 13-SEP-2005 (TREMBLrel. 31, Last sequence update)
 DE Hypothetical protein.
 GN ORFNames=232.C00006;
 OS Entamoeba histolytica HM-1:IMSS.
 CC Eukaryota; Entamoebidae; Entamoeba.
 CC NCBI_TaxID=294381;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=HM-1:IMSS;
 RX PubMed=15729342; DOI=10.1038/nature03291;
 RA Loftus B., Anderson I., Davies R., Alsmark U.C., Samuelson J.,
 RA Amedeo P., Roncaglia P., Bettiman M., Hirt R.P., Mann B.J., Nozaki T.,
 RA Suh B., Pop M., Duchene M., Ackers J., Tannich E., Leipzig M.,
 RA Hofer M., Bruchhaus I., Willhoeft U., Bhattacharya A.,
 RA Chillingworth T., Churcher C., Hance Z., Harris B., Harris D.,
 RA Jagels K., Moule S., Mungall K., Ormond D., Squares R., Whitehead S.,
 RA Quail M.A., Rabinowitz E., Norbertczak H., Price C., Wang Z.,
 RA Guillen N., Gilchrist C., Stroup S.E., Bhattacharya S., Lohia A.,
 RA Foster P.G., Sichert-Ponten T., Weber C., Singh U., Mukherjee C.,
 RA El-Sayed N.M., Petri W.A., Clark C.G., Embley T.M., Barrell B.,
 RA Frazer C.M., Hall N.,

RT "The genome of the protist parasite Entamoeba histolytica.";
 RL Nature 433:865-868(2005).
 CC -!- CAUTION: The sequence shown here is derived from an
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
 CC preliminary data.
 CC DR EMBL; AAFB01000976; EAL43455.1; -; Genomic_DNA.
 CC KW Hypothetical protein.
 SQ SEQUENCE 228 AA; 25935 MW; 95BE1A30092FD159 CRC64;

Query Match 16.7%; Score 257; DB 2; Length 228;
 Best Local Similarity 32.9%; Pred. No. 4.8e-12;
 Matches 68; Conservative 44; Mismatches 79; Indels 16; Gaps 8;

97 IQQSLSRKLVHLSGLFVLAAMPFSGSTEARPAFVPLVNGRLVINGL-SISPSNM 154
 ID LQSLSRKLVHLSGLFVLAAMPFSGSTEARPAFVPLVNGRLVINGL-SISPSNM 154
 AC OS0506_1
 DT 13-SEP-2005 (TREMBLrel. 31, Last sequence update)
 DT 13-SEP-2005 (TREMBLrel. 31, Last sequence update)
 DE Hypothetical protein.
 GN ORFNames=189.C00015;
 OS Entamoeba histolytica HM-1:IMSS.
 CC Eukaryota; Entamoebidae; Entamoeba.
 CC NCBI_TaxID=294381;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=HM-1:IMSS;
 RX PubMed=15729342; DOI=10.1038/nature03291;
 RA Loftus B., Anderson I., Davies R., Alsmark U.C., Samuelson J.,
 RA Amedeo P., Roncaglia P., Bettiman M., Hirt R.P., Mann B.J., Nozaki T.,
 RA Suh B., Pop M., Duchene M., Ackers J., Tannich E., Leipzig M.,
 RA Hofer M., Bruchhaus I., Willhoeft U., Bhattacharya A.,
 RA Chillingworth T., Churcher C., Hance Z., Harris B., Harris D.,
 RA Jagels K., Moule S., Mungall K., Ormond D., Squares R., Whitehead S.,
 RA Quail M.A., Rabinowitz E., Norbertczak H., Price C., Wang Z.,
 RA Guillen N., Gilchrist C., Stroup S.E., Bhattacharya S., Lohia A.,
 RA Foster P.G., Sichert-Ponten T., Weber C., Singh U., Mukherjee C.,
 RA El-Sayed N.M., Petri W.A., Clark C.G., Embley T.M., Barrell B.,
 RA Frazer C.M., Hall N.,

90 SITKENVIQOSTSRKLVHLSGLFVLAAMPFSGSTEARPAFVPLVNGRL 142
 ID LQSLSRKLVHLSGLFVLAAMPFSGSTEARPAFVPLVNGRLVINGL-SISPSNM 154
 AC OS0506_1
 DT 13-SEP-2005 (TREMBLrel. 31, Last sequence update)
 DT 13-SEP-2005 (TREMBLrel. 31, Last sequence update)
 DE Hypothetical protein.
 GN ORFNames=189.C00015;
 OS Entamoeba histolytica HM-1:IMSS.
 CC Eukaryota; Entamoebidae; Entamoeba.
 CC NCBI_TaxID=294381;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=HM-1:IMSS;
 RX PubMed=15729342; DOI=10.1038/nature03291;
 RA Loftus B., Anderson I., Davies R., Alsmark U.C., Samuelson J.,
 RA Amedeo P., Roncaglia P., Bettiman M., Hirt R.P., Mann B.J., Nozaki T.,
 RA Suh B., Pop M., Duchene M., Ackers J., Tannich E., Leipzig M.,
 RA Hofer M., Bruchhaus I., Willhoeft U., Bhattacharya A.,
 RA Chillingworth T., Churcher C., Hance Z., Harris B., Harris D.,
 RA Jagels K., Moule S., Mungall K., Ormond D., Squares R., Whitehead S.,
 RA Quail M.A., Rabinowitz E., Norbertczak H., Price C., Wang Z.,
 RA Guillen N., Gilchrist C., Stroup S.E., Bhattacharya S., Lohia A.,
 RA Foster P.G., Sichert-Ponten T., Weber C., Singh U., Mukherjee C.,
 RA El-Sayed N.M., Petri W.A., Clark C.G., Embley T.M., Barrell B.,
 RA Frazer C.M., Hall N.,

QY 143 VINGLISPNMILKSYTBGRABELLKGPLFYVALLFSAVFPRSPIGMISLAWMG 202
 DB 83 IKGP1A---DPLATVCRNGDYKMLTYGPLATCCIMFPLWTRNYPASTIGMILMT 138
 QY 203 GDGIADIMGRKFGSTKI--PYNPRKSWAGSISMPFIFG-----FFISIALLYYSSIGY 253
 DB 139 GDGMAEIIKGMIGKTLQKNPCKTKTIGALAVWVGAVGMWVCYIIIFGQIFIGSI-- 196
 QY 254 LHMWETTLQRVAVSMVATVYESLPITDQDDNISVPLATITLAAYLSF 302
 DB 197 -----ISGLVGAIVEFYSPYNY--DNVFIPLSSILLGFPIF 230

RESULT 14
 Q8KB08_CHLTFE PRELIMINARY; PRT; 237 AA.
 AC Q8KB08;
 DT 01-OCT-2002 (TREMBlrel. 22, Created)
 DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
 DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
 DE Hypothetical protein.
 OS OrderedLocustNames=CT1990;
 GN Chlorobium tepidum.
 OC Bacteria; Chlorobi; Chlorobia; Chlorobiales; Chlorobiaceae;
 CC Chlorobaculum.
 NCBI_TaxID=1097;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=TTLS / ATCC 49652 / DSM 12025;
 RX MEDLINE=22103685; PubMed=12093901; DOI=10.1073/pnas.132181499;
 RA Eisen J.A., Nelson K.E., Paulsen I.T., Heidelberg J.F., Wu M.,
 Raebson R.J., Deboy R.T., Gwim M.L., Nelson W.C., Holt D.H.,
 Hickey E.K., Peterson J.D., Durkin A.S., Kolonay J.F., Yang F.,
 Holt I.E., Umayam L.A., Mason T.M., Brenner M., Shea T.P.,
 Parksey D.S., Niernan W.C., Feldblum T.V., Hansen C.L., Craven M.B.,
 Radune D., Vamathevan J.J., Knouri H.M., White O., Gruber T.M.,
 Ketchum K.A., Venter J.C., Tettelin H., Bryant D.A., Fraser C.M.;
 RT "The complete genome sequence of Chlorobium tepidum TLS, a
 photoautotrophic, anaerobic, green-sulfur bacterium.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:9509-9514(2002).
 DR EMBL; AB006470; AAM73208.1; -; Genomic_DNA.
 DR TIGR; CT1990; -;
 DR GO; GO:0016020; C:membrane; IEA.
 DR GO; GO:0004605; F:phosphatidate cytidyl[yl]transferase activity; IEA.
 DR GO; GO:0008654; P:phospholipid biosynthesis; IEA.
 DR InterPro; IPR000374; PC_trans.
 DR Pfam; PF01148; CTP_transf.1; 1.
 DR Complete proteome; Hypothetical protein.
 SQ SEQUENCE 237 AA; 25945 MW; 4211f9265088BE71 CRC64;

Query Match 15.2%; Score 234; DB 2; Length 237;
 Best Local Similarity 27.2%; Pred. No. 3.1e-10;
 Matches 65; Conservative 58; Mismatches 92; Indels 24; Gaps 7;

QY 69 HDVGATVAVLGGATVAVLSPESLTGRANVIOQSLSKLVHIIISGLLFVLAWPIFGSTEAR 128
 DB 16 HNVAVMLTTIYAVSVPLMMLVYTNHGLPRDISKIRHICAGSVIYPL-PLFRDGMWSH 74
 QY 129 YFAAFVPLVNLGRLVINGLISPNMILKSYTBGRABELLKGPLFYVALLFSAVFPR 188
 DB 75 YLNTVTVAVTVLVLIQKGLFAADDQAVKTRITGDKRELKGPLFYIVAMICGTLTYK 134
 QY 189 ESPIGMISLAWMGCGDIADIMGRKFGSTKI-PYNPRKSWAGSISMPFIFG-----FFISI 242
 DB 135 QF-AGVILAMAILGWDGLAPIVGTBMGMKVKYFCERSVEGSLA-FLAGSLAAGLFFVWL 192
 QY 243 ALLYYSSLGYLHNMWETTLQRVAVSMVATVYESLPITDQDDNISVPLATI-LAAYL 300
 DB 193 IVPAQFND-----AKIMAVAAVATVIEAL--SPKRVNITLIPAEVIALAAVL 237

RESULT 15

OSOLM4_ENTHI
 ID OSOLM4_ENTHI PRELIMINARY; PRT; 230 AA.
 AC OSOLM4;
 DT 13-SEP-2005 (TREMBlrel. 31, Created)
 DT 13-SEP-2005 (TREMBlrel. 31, Last sequence update)
 DT 13-SEP-2005 (TREMBlrel. 31, Last annotation update)
 DE Hypothetical protein.
 GN ORFNames=687.t00002;
 OS Entamoeba histolytica HM-1:IMSS.
 OC Eukaryota; Entamoebidae; Entamoeba.
 NCBI_TaxID=294381;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=HM-1:IMSS;
 RX PubMed=15729342; DOI=10.1038/nature03291;
 RA Loftus B., Anderson I., Davies R., Alsmark U.C., Samuelson J.,
 Amedeo P., Roncaglia P., Berriman M., Hirt R.P., Mann B.J., Nozaki T.,
 Suh B., Pop M., Duchene M., Ackers J., Tannich E., Leippe M.,
 Hofer M., Bruchhaus I., Willhoeft U., Bhattacharya A.,
 Chillingworth T., Churcher C., Hance Z., Harris B., Harris D.,
 Jagers K., Moule S., Mungall K., Ormond D., Squares R., Whitehead S.,
 Quail M.A., Rabinowitch E., Norbertczak H., Price C., Wang Z.,
 Guillemin N., Gilchrist C., Stroup S.B., Bhattacharya S., Lohia A.,
 Foster P.G., Sichertitz-Ponten T., Weber C., Singh U., Mukherjee C.,
 El-Sayed N.M., Petri W.A., Clark C.G., Embley T.M., Barrell B.,
 Fraser C.M., Hall N.;
 RT "The genome of the eucrotist parasite Entamoeba histolytica.";
 RL Nature 433:865-868(2005).
 CC -1- CAUTION: The sequence shown here is derived from an
 EMBL/Genbank/DBJ whole genome shotgun (WGS) entry which is
 preliminary data.
 KW Hypothetical protein.
 SQ SEQUENCE 230 AA; 25851 MW; E7F78A75787B56D7 CRC64;

Query Match 15.0%; Score 230; DB 2; Length 230;
 Best Local Similarity 26.5%; Pred. No. 6.2e-10;
 Matches 67; Conservative 49; Mismatches 97; Indels 40; Gaps 7;

QY 59 SSAVATNSLHDVGATVAVLGGATVAVLSPESLTGRANVIOQSLSKLVHIIISGLLFVLA 118
 DB 9 TSLAVNVCL--IYAT-----LSKTNITITSYTSRKCVHISLGFOLLW 51
 QY 119 PIFGSTEAPRAFAFVPLVNLGRLVINGLISPNMILKSYTBGRABELLKGPLFY 175
 DB 52 KYPEEPTARIWGMCCILYAIVLIFGMKIKGVADPLATVCRNGDYKMLYGPLNY 111
 QY 176 VLLAFSAVFPWRSPIGMISLAWMGCGDIADIMGRKFGSTKI--PYNPRKSWAGSISM 233
 DB 112 CIMSFSLSLWRYRPPSIIGMAYMLTGDMAEIIKGMIGKTLQKNPCKTKTIGALAV 171
 QY 234 FIFGFFISIALLYYSSLGYLHNMWETTLQRVAVSM-----VATVYESLPITDQDDNIS 289
 DB 172 MVCC-----SLGAMFMCMLFNNFYIIESIVGFGVAIVEFYCPYNY--DNVIF 217
 QY 290 VPLATITLAAYLSF 302
 DB 218 IPLSSVWGAIFF 230

Search completed: March 16, 2006, 19:51:00
 Job time : 115 secs

THIS PAGE LEFT BLANK

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: March 16, 2006, 19:51:15 ; Search time 29 Seconds
(without alignments)
866.669 Million cell updates/sec

Title: US-10-634-548-2

Perfect score: 1536

Sequence: 1 MAATLPLSPINHQCRRGN.....DNIISVPLATILAAVLSFGY 304

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 8265679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /cgn2_6/ptodata/1/1aa/5_COMB.pep:*
- 2: /cgn2_6/ptodata/1/1aa/6_COMB.pep:*
- 3: /cgn2_6/ptodata/1/1aa/H_COMB.pep:*
- 4: /cgn2_6/ptodata/1/1aa/PTUS_COMB.pep:*
- 5: /cgn2_6/ptodata/1/1aa/R_COMB.pep:*
- 6: /cgn2_6/ptodata/1/1aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	105.5	6.9	164	US-09-248-796A-16153	Sequence 16153, A
2	100.5	6.5	285	US-09-328-352-7257	Sequence 7257, Ap
3	100.5	6.5	296	US-09-540-236-2118	Sequence 2118, Ap
4	99	6.4	515	US-09-328-352-5090	Sequence 5090, Ap
5	98.5	6.4	383	US-09-710-279-2168	Sequence 2168, Ap
6	98.5	6.4	452	US-09-134-001C-3935	Sequence 3935, Ap
7	97.5	6.3	300	US-09-543-681A-4454	Sequence 4454, Ap
8	97	6.3	431	US-09-549-848B-6	Sequence 6, Appl1
9	97	6.3	431	US-09-688-069-6	Sequence 6, Appl1
10	97	6.3	446	US-09-134-000C-6454	Sequence 6454, Ap
11	96	6.2	320	US-09-489-039A-10968	Sequence 10968, A
12	95	6.2	461	US-08-672-814D-2	Sequence 2, Appl1
13	95	6.2	461	US-09-333-696-2	Sequence 2, Appl1
14	95	6.2	461	US-09-282-218A-2	Sequence 2, Appl1
15	94.5	6.2	424	US-09-134-001C-5009	Sequence 5009, Ap
16	94	6.1	251	US-09-134-000C-4324	Sequence 4324, Ap
17	94	6.1	462	US-09-328-352-4694	Sequence 4694, Ap
18	93.5	6.1	409	US-09-710-279-2002	Sequence 2002, Ap
19	93.5	6.1	409	US-09-710-279-2306	Sequence 2306, Ap
20	93.5	6.1	414	US-09-489-039A-10869	Sequence 10869, A
21	93.5	6.1	662	US-07-841-651-4	Sequence 4, Appl1
22	93	6.1	308	US-09-328-352-7092	Sequence 7092, Ap
23	93	6.1	940	US-09-328-352-8165	Sequence 8165, Ap
24	92	6.0	538	US-09-107-532A-5563	Sequence 5563, Ap
25	91.5	6.0	542	US-09-605-703B-492	Sequence 492, App
26	91.5	6.0	688	US-09-605-703B-490	Sequence 490, App
27	91	5.9	516	US-09-949-016-11239	Sequence 11239, A

28	90	5.9	332	US-09-830-217-20	Sequence 20, Appl
29	90	5.9	332	US-10-278-946-20	Sequence 1993, Ap
30	89.5	5.8	444	US-09-540-236-1993	Sequence 1993, Ap
31	89.5	5.8	479	US-09-583-110-4687	Sequence 4687, Ap
32	89.5	5.8	480	US-09-107-433-3418	Sequence 3418, Ap
33	89	5.8	289	US-09-489-039A-12390	Sequence 12390, A
34	89	5.8	408	US-09-543-681A-4822	Sequence 4822, Ap
35	88	5.7	279	US-08-549-515-7	Sequence 7, Appl1
36	87.5	5.7	316	US-09-549-848B-33	Sequence 33, Appl
37	87.5	5.7	316	US-09-688-069-33	Sequence 33, Appl
38	87.5	5.7	348	US-09-543-681A-5845	Sequence 5845, Ap
39	87.5	5.7	397	US-09-602-787A-628	Sequence 628, App
40	87.5	5.7	629	US-09-602-787A-624	Sequence 624, App
41	87	5.7	423	US-09-328-352-6097	Sequence 6097, Ap
42	87	5.7	484	US-09-328-352-7797	Sequence 7797, Ap
43	87	5.7	544	US-09-902-540-16154	Sequence 16154, A
44	86.5	5.6	307	US-09-107-532A-6125	Sequence 6125, Ap
45	86	5.6	220	US-09-603-552-13	Sequence 13, Appl

ALIGNMENTS

```
RESULT 1
US-09-248-796A-16153
Sequence 16153, Application US/09248796A
Patent No. 6747137
GENERAL INFORMATION:
APPLICANT: Kelch Weinstock et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
FILE REFERENCE: 107196.132
CURRENT APPLICATION NUMBER: US/09/248,796A
PRIOR FILING DATE: 1999-02-12
PRIOR APPLICATION NUMBER: US 60/074,725
PRIOR FILING DATE: 1998-02-13
PRIOR APPLICATION NUMBER: US 60/096,409
NUMBER OF SEQ ID NOS: 28208
SEQ ID NO 16153
LENGTH: 164
TYPE: PRT
ORGANISM: Candida albicans
US-09-248-796A-16153

Query Match 6.9%; Score 105.5; DB 2; Length 164;
Best Local Similarity 23.2%; Pred. No. 0.0004;
Matches 38; Conservative 35; Mismatches 66; Indels 25; Gaps 8;

Qy 152 NSMLIKSVTEGRAREELKLPFYVALLFSVAFPMRESPIGMSLMMCGDGIADIMG 211
Db 2 NCSSWSPVINDNEYS-YNGLFYLGVLT--VLVAVPKDISVLSILLSNADTPAATFG 58
Qy 212 RKFGSTRIPYPR---KSWAGSISMF-----FCGFISALILYYSGLYHMMW 258
Db 59 RARPK-----YTPQASKSLAGCLASCVTGLTYLWYGF--IPQYNNWNPBEIYRNA 112
Qy 259 ETLQRYAVWSVATVRESI--PTD--QLDNIISVPLATILAAV 299
Db 113 SSNKLSPVYSILGIVTISSEVIDIADIDNFIPLSGTCIV 156

RESULT 2
US-09-328-352-7257
Sequence 7257, Application US/09328352
Patent No. 6562958
GENERAL INFORMATION:
APPLICANT: Gary L. Breton et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
FILE REFERENCE: GTC99-039A
CURRENT APPLICATION NUMBER: US/09/328,352
CURRENT FILING DATE: 1999-06-04
```

```

; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 7257
; LENGTH: 285
; TYPE: PRF
; ORGANISM: Acinetobacter baumannii
US-09-328-352-7257

```

Query Match	6.5%	Score 100.5	DB 2	Length 285
Best Local Similarity	21.2%	Pred. No. 0.0037		
Matches 48	Conservative 40	Mismatches 83	Indels 55	Gaps 8

QY 70 VVGAFVAALGGAYALVAFES---LTKNVIOQSIRKVIHLISGLLFITLAPIFGSGTE 126
 Db 61 EVGAVVKKRKMGRGILLVAFVSGVALFFHDIALMLMSASILLTVASVYWKSPFEDGMYN 120
 QY 127 ARYPAAFPVLVNGELRVINGLISPSNSMLIKSVTEGRAELLKGPLFYALLFSAYF- 185
 Db 121 ATLY-----YIGLITICAAV-----TALPV 140
 QY 186 FWRBSPICMISLAMMC-GGDGIADIMGRKFGSTKI--PYNPKKSNAGSISMTIFGFFSI 242
 Db 141 VMQSSPWWMLMYLFELLVMGADSGAYFVGGRKFGKRLAPTVSPKKSVEG-----LYGGLT 195
 QY 243 ALLVYSSSLGYLHNMWETLQRVANVSWATVYESLPIFDQDDNI 288
 Db 196 IIVMLVVOYQYINLTW---VOQLPPLISLTVFSGVADLFESMI 238

```

RESULT 3
US-09-540-236-2118
; Sequence 2118. Application US/09540236
; Patent No. 6673910
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATARRHALIS
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709 2005-001
; CURRENT APPLICATION NUMBER: US/09/540,236
; CURRENT FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 3840
; SEQ ID NO 2118
; LENGTH: 296
; TYPE: PRT
; ORGANISM: M.catarrhalis
US-09-540-236-2118

```

Query Match	6.5%	Score 100.5;	DB 2;	Length 296;
Best Local Similarity	27.0%	Pred. No. 0.004;		
Matches	40;	Conservative	28;	Mismatches 51;
				Indels 29;
				Gaps 9

```

QY      173 LEFYVALLFSAVP-----FWRESPP-IGNISLMMCGGGIADIMKRRKSGTPIPN--P 223
Db      120 LYYMGAVIITTAITANYGLMCMSPMWLVFLVWADSG-AFYIQRKSGKRRMAPNVSP 178
QY      224 RKSMAGSIMPTFGFPISTALLYYSSLG-----YLAHNMETTLQRY-----AMVSWA 272
Db      179 NKSISGLCGGLMVGSGVYTVAVGGYITQISGSLVLFILISWLTIVLASIITLGDLEFSMLKRR 238
QY      273 TVVES---LP-----ITDOLDNIS-VPL 292
Db      239 GKIDSGTILPGHGILIRIDSLISALPI 266

```

```

RESULT 4
US-09-328-352-5090
; Sequence 5090, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352

```

```

: CURRENT FILING DATE: 1999-06-04
: NUMBER OF SEQ ID NOS: 8252
: SEQ ID NO 5090
: LENGTH: 515
: TYPE: PRN
: ORGANISM: Acinetobacter baumannii
US-09-328-352-5090

```

Query Match	6.4%;	Score 99;	DB 2;	Length 515;
Best Local Similarity	24.6%;	Pred. No. 0.014;		
Matches 60;	Conservative 40;	Mismatches 94;	Indels 50;	Gaps 14;

[illegible]

```

RESULT 5
US-09-710-279-2168
; Sequence 2168, Application US/09710279
; Patent No. 6703492
; GENERAL INFORMATION:
; APPLICANT: KIMBERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PUJ480US
; CURRENT APPLICATION NUMBER: US/09/710,279
; CURRENT FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2168
; LENGTH: 383
; TYPE: PRT
; ORGANISM: Artificial Sequence
FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: amino acid sequence
US-09-710-279-2168

```

Query Match	6.4%;	Score 98.5;	DB 2;	Length 383;
Best Local Similarity	19.9%;	Pred. No. 0.01;		
Matches 67;	Conservative 49;	Mismatches 103;	Indels 117;	Gaps 13

```

Oy 13 DLCRGNNSLTHRRCSGFIILISPCFICLQMSGATOLRARRSISSAVANSIHLHDVG 72
Db 61 QIMVFKSSTLNIQET--IISNP--WLTVLGGIFILITMVLVTLGVENG--LEKXS 112
Oy 73 ATVAVLGAGAVAVLSFESLITKRNVIQ-----QSLSRKLVHLISGLFLVLAIPIS 122
Db 113 KIMMFLILFILLIIVAGSLTLEGALGVRYILLQPREVEMS-----IQGVLEAGSGFET 166
Oy 123 ---GSTEARYPAAVP-----LVNGLRLVTNGLSISFNSMLIKSVTEBGRAR 166
Db 167 LSLGTGMITVASYAPKNNITKSALSIVNNIILISVLGLAIFPALKTGYQPOEG--- 223
Oy 167 ELKKPPLFYVALLASAVPFRRRESITGMSILAMCGGCGIADIMGRKROSTILPINPRKS 226

```

Db 224 ---GGLLKVLPVSEMTF----- 240
 Qy 227 WAGSISMPFGFFSIALLYSSGLYLMNN-----ETTLQVAVNSVAVTESLP 279
 Db 241 --GTFPFYIFLLFLPAL--TSSISLELVNSFTKDNDSKQKVALIGSILVFIISIP 296
 Qy 280 IT-----DOLD--DNISVPLATI 295
 Db 297 ATLSFSSLSHLRFGAGTIFDNNDPIVSNILMPLGAL 332

RESULT 6

US-09-134-001C-3935
 ; Sequence 3935, Application US/09134001C
 ; Patent No. 6380370
 ; GENERAL INFORMATION:
 ; APPLICANT: Lynn Doucette-Stamm et al
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
 ; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: GTC-007
 ; CURRENT APPLICATION NUMBER: US/09/134.001C
 ; CURRENT FILING DATE: 1998-08-13
 ; PRIOR APPLICATION NUMBER: US 60/064,964
 ; PRIOR FILING DATE: 1997-11-08
 ; PRIOR APPLICATION NUMBER: US 60/055,779
 ; PRIOR FILING DATE: 1997-08-14
 ; NUMBER OF SEQ ID NOS: 5674
 ; SEQ ID NO 3935
 ; LENGTH: 452
 ; TYPE: PRT
 ; ORGANISM: Staphylococcus epidermidis
 ; US-09-134-001C-3935

Query Match 6.4%; Score 98.5; DB 2; Length 452;

Best Local Similarity 19.9%; Pred. No. 0.013;

Matches 67; Conservative 49; Mismatches 103; Indels 117; Gaps 13;

Qy 13 QLCRFNNSLTTHRFCSGFLISSPCFGLTGMGSATQLRARSLISSAVATNSLHDVG 72
 Db 130 QIMPFKSTLTNIQFET---IISNP---WLTVAGQGFILITMVIVMLGVEKG--LEKAS 181
 Qy 73 ATNAVGGAVYLVSESLTKRNVQ-----QSLSRKLVIIISGLIFVLAMPFIS 122
 Db 182 KIMPELFIPLITVAQSLTLEGAGVRYLIQPREVMS-----IQGVFALGQSFT 235
 Qy 123 ---GSTEARYPAAPV-----LVNGLRLVINGLSISPSMLIKSVTEGRAB 166
 Db 236 LSLGTGMITAYASVPRKMTIKSSALSIYVNNILSVLAGLAIFALTKTFGQPEG--- 292
 Qy 167 ELKGPLFYVALLFSAPVPMRESPIGMSILAMCGGDGIADIMGRKFGSTKIPYPRKS 226
 Db 293 ---GGLLKVLPVSEMTF----- 309
 Qy 227 WAGSISMPFGFFSIALLYSSGLYLMNN-----ETTLQVAVNSVAVTESLP 279
 Db 310 --GTFPFYIFLLFLPAL--TSSISLELVNSFTKDNDSKQKVALIGSILVFIISIP 365
 Qy 280 IT-----DOLD--DNISVPLATI 295
 Db 366 ATLSFSSLSHLRFGAGTIFDNNDPIVSNILMPLGAL 401

RESULT 7

US-09-543-681A-4454
 ; Sequence 4454, Application US/09543681A
 ; Patent No. 6605709
 ; GENERAL INFORMATION:
 ; APPLICANT: GARY BRETON
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
 ; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 2709.1002-001
 ; CURRENT APPLICATION NUMBER: US/09/543,681A

; CURRENT FILING DATE: 2000-04-05
 ; PRIOR APPLICATION NUMBER: US 60/128,706
 ; PRIOR FILING DATE: 1999-04-09
 ; NUMBER OF SEQ ID NOS: 8344
 ; SEQ ID NO 4454
 ; LENGTH: 300
 ; TYPE: PRT
 ; ORGANISM: Proteus mirabilis
 ; US-09-543-681A-4454

Query Match

6.3%; Score 97.5; DB 2; Length 300;

Best Local Similarity 24.9%; Pred. No. 0.0092;

Matches 66; Conservative 38; Mismatches 78; Indels 83; Gaps 17;

Qy 31 GFLISSPCFGLTGMGSATQLRARSLISSAVATNSLHDV--GATVAVGGAVYLVLSF 88
 Db 72 GFLIS-----MHHG--TPMRARNSF--RFGHSSLSNEIASGATFPALGGLTWLI--- 118
 Qy 89 ESLTRNVIIQSLRKVHIISGLIFVLA-----WPIFGSTEARY--PAAPV- 134
 Db 119 SVLNKMPALCKLWLVVMILA-VLFTAIRVYQIDVPTW--YNGITVNFVLTARIG 175
 Qy 135 -PLVNGRLVINGLSISPSMLIKSVTEGRABELKGPLFYVALLFSAPVPMRES--PI 192
 Db 176 GPILAALIMRIAGFTNCTSL-----PLLSVIAIITISALVATSGQREL 219
 Qy 193 GMISLAMCGDGIAD--IMGRKFGSTKI-----PYNPKSWAGSISMPFICGF 238
 Db 220 GSIOYSQKAVDLPVDYGFELMGIILALITGLACMIAPLARKNP-----SVSILMVIGF 273
 Qy 239 -----FISIALYSSGLYLMH 256
 Db 274 ILVFGVEFTGRGVF-----GLHM 292

RESULT 8

US-09-549-848B-6

; Sequence 6, Application US/09549848B

; Patent No. 6541259

; GENERAL INFORMATION:

; APPLICANT: Laesner, Michael

; APPLICANT: Post-Bettenmiller, Dusty

; APPLICANT: Savidge, Beth

; TITLE OF INVENTION: Nucleic Acid Sequences Involved in

; FILE REFERENCE: 17133/02/US

; CURRENT APPLICATION NUMBER: US/09/549,848B

; CURRENT FILING DATE: 2000-04-14

; PRIOR APPLICATION NUMBER: 60/129,899

; PRIOR FILING DATE: 1999-04-15

; PRIOR APPLICATION NUMBER: 60/146,461

; PRIOR FILING DATE: 1999-07-30

; NUMBER OF SEQ ID NOS: 94

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 6

; LENGTH: 431

; TYPE: PRT

; ORGANISM: Arabidopsis sp

; US-09-549-848B-6

Query Match 6.3%; Score 97; DB 2; Length 431;

Best Local Similarity 21.9%; Pred. No. 0.019;

Matches 76; Conservative 50; Mismatches 121; Indels 100; Gaps 18;

Qy 1 MAATLP--LSPIHOLCFRNNLSLTHRFCSGFLISSPCFGLTGMGS--ATQLRARS 56
 Db 16 VSSILNPRLLPMSRELC-----AVNSFQPFVSTSTAKDITGVRSDANVFATAT 68
 Qy 57 LISSAVATNSLHDVGATVAVLGAYA-----LVLS 87
 Db 69 AATATATGTGSSRVALAGLGHYVACWELSKAKLSMLVAVTSGTGTYLGTGNAIS 128

Db 256 IVPBMRGADPKRPEKRAIVGMAITPTLLALVPLSVISLNGLDNIDSVATISWGRALG 315

Cy 216 STKIPNPKSMAGS 230

Db 316 -----EMASS 320

RESULT 12

US-08-672-814D-2
Sequence 2, Application US/08672814D

Patent No. 5952480

GENERAL INFORMATION:

APPLICANT: Leung, David W.

TITLE OF INVENTION: MAMMALIAN

NUMBER OF SEQUENCES: 13

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Cell Therapeutics, Inc.

STREET: 201 Elliott Avenue West

CITY: Seattle

STATE: Washington

COUNTRY: U.S.A.

ZIP: 98119

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5" disk, 1.44MB, double side, high density

COMPUTER: PC Clone (486 microprocessor)

OPERATING SYSTEM: MS-DOS Version 6.1, Windows NT

SOFTWARE: WORD 6.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/672,814D

FILING DATE: 28-Jun-1996

ATTORNEY/AGENT INFORMATION:

NAME: Paciszewski, Stephen

REGISTRATION NUMBER: 36,131

REFERENCE/DOCKET NUMBER: 1803

TELECOMMUNICATION INFORMATION:

TELEPHONE: (206)282-7100

TELEFAX: (206)284-6206

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 461

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: polypeptide

HYPOTHETICAL: no

ANTI-SENSE: NO

FRAGMENT TYPE:

ORIGINAL SOURCE:

ORGANISM: homo sapien

STRAIN:

INDIVIDUAL ISOLATE:

DEVELOPMENTAL STAGE:

HAPLOTYPE:

TISSUE TYPE:

CELL TYPE:

ORGANELLE:

FEATURE:

NAME/KEY: CDS

LOCATION:

IDENTIFICATION METHOD:

OTHER INFORMATION:

US-08-672-814D-2

Query Match 6.2%; Score 95; DB 1; Length 461;

Best Local Similarity 26.9%; Pred. No. 0.036; Matches 42; Conservative 20; Mismatches 54; Indels 40; Gaps 8;

Db 121 FSGSTEARYPFAFP-----LVNGLRVINGLSISPSNMLIKSVTEGRABELLKGPL 173

Db 157 FYGSETVADYFATFQREBQQLIRYHRFISFALYLAFCMFVSLVGEHRLQ----- 210

Cy 174 PYVLA-----LLFSAVF---FYRESPIGMSIAMCGGDIADIMGRKGS 216

Db 211 FYMFAMWHTVLLITVTSQHLVQLNLFEGMTWFLVPIS---SVIC-NDITAYLFGFFGR 265

Cy 217 TK-IPNPKSMAGSISMF---IFGFFISALLYY 247

Db 266 TPLKLSPKTWEGFISGFFSTVVFGLIAVVLISKY 301

RESULT 13

US-09-333-696-2

Sequence 2, Application US/09333696

Patent No. 6200769

GENERAL INFORMATION:

APPLICANT: Leung, David W.

TITLE OF INVENTION: MAMMALIAN

NUMBER OF SEQUENCES: 13

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Cell Therapeutics, Inc.

STREET: 201 Elliott Avenue West

CITY: Seattle

STATE: Washington

COUNTRY: U.S.A.

ZIP: 98119

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5" disk, 1.44MB, double side, high density

COMPUTER: PC Clone (486 microprocessor)

OPERATING SYSTEM: MS-DOS Version 6.1, Windows NT

SOFTWARE: WORD 6.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/333,696

FILING DATE:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/672,814

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Paciszewski, Stephen

REGISTRATION NUMBER: 36,131

REFERENCE/DOCKET NUMBER: 1803

TELECOMMUNICATION INFORMATION:

TELEPHONE: (206)282-7100

TELEFAX: (206)284-6206

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 461

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: polypeptide

HYPOTHETICAL: no

ANTI-SENSE: NO

FRAGMENT TYPE:

ORIGINAL SOURCE:

ORGANISM: homo sapien

STRAIN:

INDIVIDUAL ISOLATE:

DEVELOPMENTAL STAGE:

HAPLOTYPE:

TISSUE TYPE:

CELL TYPE:

ORGANELLE:

FEATURE:

NAME/KEY: CDS

LOCATION:

IDENTIFICATION METHOD:

OTHER INFORMATION:

US-09-333-696-2

Query Match 6.2%; Score 95; DB 2; Length 461;

Best Local Similarity 26.9%; Pred. No. 0.036;

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: March 16, 2006, 19:51:40 ; Search time 79 Seconds
(without alignments)
1607.850 Million cell updates/sec

Title: US-10-634-548-2
Perfect score: 1536
Sequence: 1 MAATLPLSPINHQLCRFGNN.....DNISVPLATLAAVLSRGY 304

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA Main:
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/US11_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1536	100.0	304	US-10-634-548-2	Sequence 2, Appl1
2	1536	100.0	315	US-10-425-114-54545	Sequence 54545, A
3	1523	99.2	301	US-10-425-114-54546	Sequence 54546, A
4	1266	82.4	304	US-10-634-548-39	Sequence 39, Appl1
5	1266	82.4	319	US-10-425-114-54630	Sequence 54630, A
6	1140	74.2	226	US-10-634-548-4	Sequence 4, Appl1
7	870	56.6	306	US-10-425-114-54635	Sequence 54635, A
8	864	56.2	298	US-10-634-548-41	Sequence 41, Appl1
9	857.5	55.8	302	US-10-634-548-45	Sequence 45, Appl1
10	791.5	51.5	314	US-10-634-548-46	Sequence 46, Appl1
11	791.5	51.5	332	US-10-425-114-55785	Sequence 55785, A
12	791.5	51.5	657	US-10-634-548-49	Sequence 49, Appl1
13	791.5	51.5	657	US-10-437-963-179783	Sequence 179783, A
14	791.5	51.5	803	US-10-634-548-48	Sequence 48, Appl1
15	791.5	51.5	803	US-10-437-963-179784	Sequence 179784, A
16	786	51.2	309	US-10-425-114-55436	Sequence 55436, A
17	782.5	50.9	269	US-10-634-548-47	Sequence 47, Appl1
18	777	50.6	303	US-10-634-548-61	Sequence 61, Appl1
19	776	50.5	251	US-10-425-114-55867	Sequence 55867, A
20	776	50.5	268	US-10-425-115-303416	Sequence 303416, A
21	773	50.3	300	US-10-634-548-58	Sequence 58, Appl1
22	773	50.3	312	US-10-425-114-55641	Sequence 55641, A
23	752.5	49.0	267	US-10-634-548-62	Sequence 62, Appl1
24	751	48.9	233	US-10-425-114-55631	Sequence 55631, A
25	751	48.9	236	US-10-425-114-55714	Sequence 55714, A
26	751	48.9	236	US-10-634-548-63	Sequence 63, Appl1
27	718.5	46.8	292	US-10-634-548-44	Sequence 44, Appl1

28	646	42.1	199	4	US-10-425-114-72714	Sequence 72714, A
29	615	40.0	191	4	US-10-634-548-55	Sequence 55, Appl1
30	615	40.0	191	4	US-10-767-701-42774	Sequence 42774, A
31	607	39.5	259	5	US-10-739-930-8545	Sequence 8545, Ap
32	604	39.3	265	4	US-10-424-599-255091	Sequence 255091,
33	591	38.5	188	4	US-10-634-548-60	Sequence 60, Appl1
34	583	38.0	211	4	US-10-425-115-303417	Sequence 303417,
35	580.5	37.8	292	4	US-10-767-701-43330	Sequence 43330, A
36	572.5	37.3	312	4	US-10-634-548-65	Sequence 65, Appl1
37	572.5	37.3	313	4	US-10-425-115-357166	Sequence 357166,
38	572.5	37.3	338	4	US-10-425-114-55756	Sequence 55756, A
39	565.5	36.8	246	4	US-10-634-548-50	Sequence 50, Appl1
40	562.5	36.6	302	4	US-10-634-548-54	Sequence 54, Appl1
41	552.5	36.0	296	4	US-10-425-114-54640	Sequence 54640, A
42	552.5	36.0	296	4	US-10-634-548-56	Sequence 56, Appl1
43	544	35.4	346	5	US-10-739-930-10410	Sequence 10410, A
44	543.5	35.4	319	4	US-10-424-599-213867	Sequence 213867,
45	542.5	35.3	319	4	US-10-634-548-43	Sequence 43, Appl1

ALIGNMENTS

RESULT 1

US-10-634-548-2 Application US/10634548

Publication No. US20040045051A1

GENERAL INFORMATION:

APPLICANT: No. US20040045051A1ris, Susan R

APPLICANT: Lincoln, Kim

APPLICANT: Abad, Mark Scott

APPLICANT: Eilers, Robert

APPLICANT: Hartsuyker, Karen Kindle

APPLICANT: Hirschberg, Joseph

APPLICANT: Karunanandaa, Balasubramanian

APPLICANT: Moshiri, Farhad

APPLICANT: Stein, Joshua C.

APPLICANT: Valentini, Henry B.

APPLICANT: Venkatesh, Tyamagondlu V.

TITLE OF INVENTION: Tocopherol biosynthesis related genes and used thereof

FILE REFERENCE: Ren-01-125

CURRENT APPLICATION NUMBER: US/10/634, 548

CURRENT FILING DATE: 2003-08-05

PRIOR APPLICATION NUMBER: us 60/400, 689

PRIOR FILING DATE: 2002-08-05

NUMBER OF SEQ ID NOS: 79

SOFTWARE: PatentIn version 3.2

SEQ ID NO 2

LENGTH: 304

TYPE: PRT

ORGANISM: Arabidopsis thaliana

US-10-634-548-2

Query Match

Best Local Similarity 100.0%; Pred. No. 11e-151;

Matches 304; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MAATLPLSPINHQLCRFGNNSLTTRFCSPGFLISSPCFGLTGMGATQARARSLISS	60
DB	1	MAATLPLSPINHQLCRFGNNSLTTRFCSPGFLISSPCFGLTGMGATQARARSLISS	60
QY	61	AAATNSLTHDVGAVAVLGAVALVLSPESTTKRNVQOQISRLVHLISGLFLVLA MPI	120
DB	61	AAATNSLTHDVGAVAVLGAVALVLSPESTTKRNVQOQISRLVHLISGLFLVLA MPI	120
QY	121	FSGSTEARFPAFVPLVNGRLVINGSLSPNSMLISVTRGAEELKGPFLVYVALL	180
DB	121	FSGSTEARFPAFVPLVNGRLVINGSLSPNSMLISVTRGAEELKGPFLVYVALL	180
QY	181	FSAVFPRESPIGMSIAMCGGDIADIMGRKFGSTKIPYPRKXWAGSISMTFGPFI	240
DB	181	FSAVFPRESPIGMSIAMCGGDIADIMGRKFGSTKIPYPRKXWAGSISMTFGPFI	240

Qy	241	SIALLYYSSLGYLHNHMETLORVAMSMATVYESLPITQDLDNISVPLATIIAAYL	3008
Db	241	SIALLYYSSLGYLHNHMETLORVAMSMATVYESLPITQDLDNISVPLATIIAAYL	3008
Qy	301	SPFGY 304	
Db	301	SPFGY 304	

RESULT 2
US-10-425-114-54545

```

Sequence 54545, Application US/10425114
Publication No. US20040034888A1
GENERAL INFORMATION:
APPLICANT: Liu, Jingtong
APPLICANT: Zhou, Yihua
APPLICANT: Kovalic, David K.
APPLICANT: Screen, Steven E
APPLICANT: Tabaska, Jack E
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(53313)B
CURRENT APPLICATION NUMBER: US/10/425,114
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 73128
SEQ ID NO 54545
LENGTH: 315
TYPE: prt
ORGANISM: Arabidopsis thaliana columbia
FEATURE:
OTHER INFORMATION: Clone ID: LIB3278-077-D5_F1.1.pep
US-10-425-114-54545

```

Query Match	100.0%	Score 1536;	DB 4;	Length 315;
Best Local Similarity	100.0%	Pred. No. 1.2e-151;		
Matches 304;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0

Qy	1	MAATLPLSGINHLQOLRFGNNSLTTHRFCSGPGFLISSPCFGLTGMASATOLRRRRSLISS	60
Db	12	MAATPLPSLFINQLCFGNNSLTTHRFCSGPGFLISSPCFGLTGMASATOLRRRRSLISS	71
Qy	61	AVATNSLHDVCAATVAVLGCAVALYVFSFSLTKRNTIQQSLSRKLVHIIISGLLFVLAWP1	120
Db	72	AVATNSLHDVCAATVAVLGCAVALYVFSFSLTKRNTIQQSLSRKLVHIIISGLLFVLAWP1	131
Qy	121	FSGSTEARYPFAFVPLVNLGLRLVINGLSISSPNSMLIKSVTRREGRAEELIKGPLEFVYLAAL	180
Db	132	FSGSTEARYPFAFVPLVNLGLRLVINGLSISSPNSMLIKSVTRREGRAEELIKGPLEFVYLAAL	191
Qy	161	FSAVFPMWRSPIGMISLAMCGGDGIADIMGRKFGSTKLPYPNRKSMAGISIMFIFGFFI	240
Db	192	FSAVFPMWRSPIGMISLAMCGGDGIADIMGRKFGSTKLPYPNRKSMAGISIMFIFGFFI	251
Qy	241	SIALLIYYSSLGYLHNHWETTILQRYVAMSVNAIVVSSLPITTDOLDNISVPLATITIAAYL	300
Db	252	SIALLIYYSSLGYLHNHWETTILQRYVAMSVNAIVVSSLPITTDOLDNISVPLATITIAAYL	311
Qy	301	SFGY 304	
Db	312	SFGY 315	

```

3
RESULT:3
US-10-425-114-54546
; Sequence 54546, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingtong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaka, Jack B

```

```

? APPLICANT: Cao, Yongwei
? TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
? TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
? FILE REFERENCE: 38-21(5313)B
? CURRENT APPLICATION NUMBER: US/10/425,114
? CURRENT FILING DATE: 2003-04-28
? NUMBER OF SEQ ID NOS: 73128
? SEQ ID NO 54546
? LENGTH: 301
? TYPE: PRT
? ORGANISM: Arabidopsis thaliana
? FEATURE:
? OTHER INFORMATION: Clone ID: LIB23-048-A2_F11.pep
US-10-425-114-54546

```

Query Match	99.2%	Score 1523;	DB 4;	Length 301;
Best Local Similarity	100.0%	Pred. No. 2.6e-150;		
Matches 301; Conservative	0;	Mismatches 0;	Indels 0;	Gaps 0;

QY	4	LLPSPINHOLCRFENNBSLITTHRFCSPPFLISSPCFII GLTGGSATQTLARBSLSSAVA	63
Db	1	TLPLSPINHOLCRFENNBSLITTHRFCSPPFLISSPCFII GLTGGSATQTLARBSLSSAVA	60
QY	64	TNSLHHDGATVAVTGGAVALVSTESLTCKNAVTOOSLRKLVHLSGLPFLVAMPISFG	123
Db	61	TNSLHHDGATVAVTGGAVALVSTESLTCKNAVTOOSLRKLVHLSGLPFLVAMPISFG	120
QY	124	STEARYFAAPFLVNGLRVLVINGLSISPSNMLIKSVTEGRAEBLLKGPLFVYALLFSA	183
Db	121	STEARYFAAPFLVNGLRVLVINGLSISPSNMLIKSVTEGRAEBLLKGPLFVYALLFSA	180
QY	184	VFPWFRESITGMISLMMCGDGIADIMGRKKESTLIPNPKRSMGSIIMPFIFGEPISTA	243
Db	181	VFPWFRESITGMISLMMCGDGIADIMGRKKESTLIPNPKRSMGSIIMPFIFGEPISTA	240
QY	244	LLYYSSSGLYHMMNETTLORVAVSMVATVVESLPITDQLDNDNISVPLATIIIAAYLSFG	303
Db	241	LLYYSSSGLYHMMNETTLORVAVSMVATVVESLPITDQLDNDNISVPLATIIIAAYLSFG	300
QY	304	Y 304	
Db	301	Y 301	

```

RESULT 4
US-10-634-548-39
; Sequence 39, Application US/10634548
; Publication No. US20040045051A1
; GENERAL INFORMATION:
; APPLICANT: No. US20040045051A1, Susan R
; APPLICANT: Lincoln, Kim
; APPLICANT: Abad, Mark Scott
; APPLICANT: Eilers, Robert
; APPLICANT: Hartsuyker, Karen
; APPLICANT: Hirschberg, Joseph
; APPLICANT: Karunanandaa, Balasubojini
; APPLICANT: Moshiri, Farhad
; APPLICANT: Stein, Joshua C.
; APPLICANT: Valentin, Henry E.
; APPLICANT: Venkatesh, Tyamagondlu V.
; TITLE OF INVENTION: Tocopherol biosynthesis related genes and used thereof
; FILE REFERENCE: Ran-01-125
; CURRENT APPLICATION NUMBER: US/10/634,548
; CURRENT FILING DATE: 2003-08-05
; PRIOR APPLICATION NUMBER: us 60/400,689
; PRIOR FILING DATE: 2002-08-05
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 39
; LENGTH: 304
; TYPE: PRT
; ORGANISM: Brassica napus
US-10-634-548-39

```

Query Match 82.4%; Score 1266; DB 4; Length 304;

Best Local Similarity 81.4%; Pred. No. 1.9e-123; Indels 6; Gaps 2;
Matches 250; Conservative 24; Mismatches 27;

```

Qy 1 MAATLPLSPINHQCIRFGNN--SLTTHRFCSPGFLISSPCFGLTGMGSAATOLRARRSL 57
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 1 MAALPLSPVSHQCIRSNRFRWYNAAMTRPCSP---VSSPCYIGVKIGSSSQARARHPL 57
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 58 ISSAVATNSLHDVGATVAALVGAYALVLSFESLTKRNVIOQSLSRKLVIHLSGLLFVLA 117
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 58 ISSASTDYLDHVGATVAALVGAYALVLSFESLTKRNVIOQRLSRKLVHLSGLLFALS 117
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 118 WPTSGSTEARYPFAFVPLVNGRLVINGLSISPSNMLIKSVTREGRAEELKGPLFVYL 177
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 118 WPTSGSTEARYPFAFVPLVNGRLVINGLSVSPSTLIQSVTRGREGREELKGPLFVYL 177
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 178 ALPSAVFWRRESPIGMSIPLAMCGGDGIADIMGRKFGSTKIPYNPRKSMAGSISMFTFG 237
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 178 ALVAAVFWRMDSPFGMISLAMCGGDGIADIMGRKXGKIPYNPRKSLAGSISMFTFG 237
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 238 PFISIALYYSSSLGYLHMNETTLQRYAMVSMVATVESLPTDQDDNISVPLATITLA 297
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 238 PFISIGLLYYSSSLGYLHMNETTFTRVAIVSLVATLVESLPTDQDDNVSVELATITLA 297
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 298 AYLSFGY 304
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 298 AYLSFGY 304
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

```

RESULT 5

US-10-425-114-54630

Sequence 54630, Application US/10425114

Publication No. US20040034888A1

GENERAL INFORMATION:

APPLICANT: Liu, Jindong

APPLICANT: Zhou, Yihua

APPLICANT: Kovalic, David K.

APPLICANT: Screen, Steven E

APPLICANT: Tabaska, Jack E

APPLICANT: Cao, Yongwei

TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

FILE REFERENCE: 38-21(5313)B

CURRENT APPLICATION NUMBER: US/10/425,114

CURRENT FILING DATE: 2003-04-28

NUMBER OF SEQ ID NOS: 73128

SEQ ID NO 54630

LENGTH: 319

TYPE: PRT

ORGANISM: Brassica napus

FEATURE:

OTHER INFORMATION: Clone ID: LIB82-020-B9_P11.pap

US-10-425-114-54630

Query Match 82.4%; Score 1266; DB 4; Length 319;

Best Local Similarity 81.4%; Pred. No. 2.1e-123;

Matches 250; Conservative 24; Mismatches 27; Indels 6; Gaps 2;

```

Qy 1 MAATLPLSPINHQCIRFGNN--SLTTHRFCSPGFLISSPCFGLTGMGSAATOLRARRSL 57
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 16 MAALPLSPVSHQCIRSNRFRWYNAAMTRPCSP---VSSPCYIGVKIGSSSQARARHPL 72
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 58 ISSAVATNSLHDVGATVAALVGAYALVLSFESLTKRNVIOQSLSRKLVIHLSGLLFVLA 117
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 73 ISSASTDYLDHVGATVAALVGAYALVLSFESLTKRNVIOQRLSRKLVHLSGLLFALS 132
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 118 WPTSGSTEARYPFAFVPLVNGRLVINGLSISPSNMLIKSVTREGRAEELKGPLFVYL 177
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 133 WPTSGSTEARYPFAFVPLVNGRLVINGLSVSPSTLIQSVTRGREGREELKGPLFVYL 192
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 178 ALPSAVFWRRESPIGMSIPLAMCGGDGIADIMGRKFGSTKIPYNPRKSMAGSISMFTFG 237
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

```

Db 193 ALVAAVFWRMDSPFGMISLAMCGGDGIADIMGRKXGKIPYNPRKSLAGSISMFTFG 252

Qy 238 PFISIALYYSSSLGYLHMNETTLQRYAMVSMVATVESLPTDQDDNISVPLATITLA 297

Db 253 PFISIGLLYYSSSLGYLHMNETTFTRVAIVSLVATLVESLPTDQDDNVSVELATITLA 312

Qy 298 AYLSFGY 304

Db 313 AYLSFGY 319

RESULT 6

US-10-634-548-4

Sequence 4, Application US/10634548

Publication No. US20040045051A1

GENERAL INFORMATION:

APPLICANT: No. US20040045051A1, Susan R

APPLICANT: Lincoln, Kim

APPLICANT: Abad, Mark Scott

APPLICANT: Elser, Robert

APPLICANT: Hartsuyker, Karen Kindle

APPLICANT: Hirschberg, Joseph

APPLICANT: Karunanandaa, Balasubramanian

APPLICANT: Moshiri, Farhad

APPLICANT: Stein, Joshua C.

APPLICANT: Valentini, Henry E.

APPLICANT: Venkatesh, Tyamagondlu V.

TITLE OF INVENTION: Tocopherol biosynthesis related genes and used thereof

FILE REFERENCE: Ren-01-125

CURRENT APPLICATION NUMBER: US/10/634,548

CURRENT FILING DATE: 2003-08-05

PRIOR APPLICATION NUMBER: us 60/400,689

PRIOR FILING DATE: 2002-08-05

NUMBER OF SEQ ID NOS: 79

SOFTWARE: PatentIn version 3.2

SEQ ID NO 4

LENGTH: 226

TYPE: PRT

ORGANISM: Arabidopsis thaliana

US-10-634-548-4

Query Match 74.2%; Score 1140; DB 4; Length 226;

Best Local Similarity 100.0%; Pred. No. 1.9e-110;

Matches 226; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAATLPLSPINHQCIRFGNN--SLTTHRFCSPGFLISSPCFGLTGMGSAATOLRARRSL 60

Db 1 MAATLPLSPINHQCIRFGNN--SLTTHRFCSPGFLISSPCFGLTGMGSAATOLRARRSL 60

Qy 61 AVATNSLHDVGATVAALVGAYALVLSFESLTKRNVIOQSLSRKLVIHLSGLLFVLA 120

Db 61 AVATNSLHDVGATVAALVGAYALVLSFESLTKRNVIOQSLSRKLVIHLSGLLFVLA 120

Qy 121 FSGSTEARYPFAFVPLVNGRLVINGLSISPSNMLIKSVTREGRAEELKGPLFVYL 180

Db 121 FSGSTEARYPFAFVPLVNGRLVINGLSISPSNMLIKSVTREGRAEELKGPLFVYL 180

Qy 181 FSAVFWRESPIGMSIPLAMCGGDGIADIMGRKFGSTKIPYNPRKS 226

Db 181 FSAVFWRESPIGMSIPLAMCGGDGIADIMGRKFGSTKIPYNPRKS 226

RESULT 7

US-10-425-114-54635

Sequence 54635, Application US/10425114

Publication No. US20040034888A1

GENERAL INFORMATION:

APPLICANT: Liu, Jindong

APPLICANT: Zhou, Yihua

APPLICANT: Kovalic, David K.

APPLICANT: Screen, Steven E

APPLICANT: Tabaska, Jack E

APPLICANT: Cao, Yongwei

TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
 TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
 FILE REFERENCE: 38-21(5313)B
 CURRENT APPLICATION NUMBER: US/10/425,114
 CURRENT FILING DATE: 2003-04-28
 NUMBER OF SEQ ID NOS: 73128
 SEQ ID NO 54635
 LENGTH: 306
 TYPE: PRT
 ORGANISM: Gossypium hirsutum
 FEATURE:
 OTHER INFORMATION: Clone ID: LJB3196-048-A11_F11.pep
 US-10-425-114-54635

Query Match 56.6%; Score 870; DB 4; Length 306;
 Best Local Similarity 56.7%; Pred. No. 4,9e-82;
 Matches 174; Conservative 51; Mismatches 72; Indels 10; Gaps 2;

QY 2 AATPLSPINHQCFGNNSLTTRFCS-----PGFLISSPCFGLTGMGSAATOLARRSL 57
 DB 6 AATNLS-----LSFTPIILSRHYSAVFPFPRFLFSLPIPTSRPILYRAPQRATA 59
 QY 58 ISSAAVATNSLHDVATVAVLGGAYALVSFESLTGRNVIOQSLSRKLVIHLSGLPYLA 117
 DB 60 LSATVATSIERDTRMASAVFAGAYALVFTFDILTOKELIQONLSRKLVIHLSGLFAIS 119
 QY 118 WPIPGSTEARYPFAFVPLVNGLRVINGLSISPSNMLIKSVTRGRABELLKGPLPYVL 177
 DB 120 WPIFNADEARYPFALVPLFNCRLVHLSLTDQSLIKSVTRGRNPKELRGPLYVA 179
 QY 178 ALFSAVFWRBSPICMISLAMCGDGIADIIMGRKFGSTKIPYPRKSMAGSISMPFEG 237
 DB 180 MMLCALVWRBSPICVCLAMCGDGVADIIGKYSKIPYNGSISWGISIMFSG 239
 QY 238 PFISIALIYYSSGLYHNMETTLQRYAMVSMVATVYESLPITDLDNISVPLATILA 297
 DB 240 FIIISGMYYSALGYLQDMGYTLHRAVATVATVYESLPISMLIDNISVPLASMLA 299
 QY 298 AYLISFGY 304
 DB 300 AYLIRFGH 306

RESULT 8

US-10-634-548-41
 Sequence 41, Application US/10634548
 Publication No. US20040045051A1
 GENERAL INFORMATION:
 APPLICANT: No. US20040045051A1ris, Susan R
 APPLICANT: Lincoln, Kim
 APPLICANT: Abad, Mark Scott
 APPLICANT: Eilers, Robert
 APPLICANT: Hartsuyker, Karen Kindle
 APPLICANT: Hirschberg, Joseph
 APPLICANT: Karunananda, Balasulojini
 APPLICANT: Moshiri, Farhad
 APPLICANT: Stein, Joshua C.
 APPLICANT: Valentin, Henry E.
 APPLICANT: Venkatesh, Tyamagondlu V.
 TITLE OF INVENTION: Tocopherol biosynthesis related genes and used thereof
 FILE REFERENCE: Ren-01-125
 CURRENT APPLICATION NUMBER: US/10/634,548
 CURRENT FILING DATE: 2003-08-05
 PRIOR APPLICATION NUMBER: us 60/400,689
 PRIOR FILING DATE: 2002-08-05
 NUMBER OF SEQ ID NOS: 79
 SOFTWARE: PatentIn version 3.2
 SEQ ID NO 41
 LENGTH: 298
 TYPE: PRT
 ORGANISM: Gossypium hirsutum
 US-10-634-548-41

Query Match 56.2%; Score 864; DB 4; Length 298;
 Best Local Similarity 60.1%; Pred. No. 2e-81;
 Matches 167; Conservative 46; Mismatches 65; Indels 0; Gaps 0;

QY 27 PCSPGLISSPCFGLTGMGSAATOLARRSLISSAVATNSLHDVATVAVLGGAYALVL 86
 DB 21 FPPPRFLFSLPIPTSRFPILYRAPQRAVALSATVATASIFRTAASAVFAGAYALVF 80
 QY 87 SFESLTGRNVIOQSLSRKLVIHLSGLFLVAMPFSGSTEARYPFAFVPLVNGLRVING 146
 DB 81 TFDILTOKELIQONLSRKLVIHLSGLFLVAMPFSGSTEARYPFAFVPLVNGLRVING 140
 QY 147 LSISPSNMLIKSVTRGRABELLKGPLPYVALIFSAVFPWRBSPICMISLAMCGDGI 206
 DB 141 LSLTDQSLIKSVTRGRNPKELRGPLYVAMMLCALVWRBSPICVCLAMCGDGV 200
 QY 207 ADIMGRKFGSTKIPYPRKSMAGSISMPFEGFISIALIYYSSGLYHNMETTLQRYA 266
 DB 201 ADIIGRKYGSKIPYNGSKSMWGISIMFSGFIIISGMYYSALGYLQDMGYTLHRAVA 260
 QY 267 MVMWATVYESLPITDLDNISVPLATILAAYLSFGY 304
 DB 261 FIVLATVYESLPISMLIDNISVPLASMLAAYLIRFGH 298

RESULT 9

US-10-634-548-45
 Sequence 45, Application US/10634548
 Publication No. US20040045051A1
 GENERAL INFORMATION:
 APPLICANT: No. US20040045051A1ris, Susan R
 APPLICANT: Lincoln, Kim
 APPLICANT: Abad, Mark Scott
 APPLICANT: Eilers, Robert
 APPLICANT: Hartsuyker, Karen Kindle
 APPLICANT: Hirschberg, Joseph
 APPLICANT: Karunananda, Balasulojini
 APPLICANT: Moshiri, Farhad
 APPLICANT: Stein, Joshua C.
 APPLICANT: Valentin, Henry E.
 APPLICANT: Venkatesh, Tyamagondlu V.
 TITLE OF INVENTION: Tocopherol biosynthesis related genes and used thereof
 FILE REFERENCE: Ren-01-125
 CURRENT APPLICATION NUMBER: US/10/634,548
 CURRENT FILING DATE: 2003-08-05
 PRIOR APPLICATION NUMBER: us 60/400,689
 PRIOR FILING DATE: 2002-08-05
 NUMBER OF SEQ ID NOS: 79
 SOFTWARE: PatentIn version 3.2
 SEQ ID NO 45
 LENGTH: 302
 TYPE: PRT
 ORGANISM: Glycine max
 US-10-634-548-45

Query Match 55.8%; Score 857.5; DB 4; Length 302;
 Best Local Similarity 60.6%; Pred. No. 9.8e-81;
 Matches 175; Conservative 44; Mismatches 51; Indels 19; Gaps 5;

QY 19 NNSLTTR-RF-CSPGLISSPCFGLTGMGSAATOL--RARRSLISSAVATNSLHDVGA 73
 DB 26 NSPTNTHVRFLCSP-----GVPPAVRLDQRLRFPVPPGAGA-EDLLYNAGA 71
 QY 74 TVAVTGGAYALVSFESLTGRNVIOQSLSRKLVIHLSGLFLVAMPFSGSTEARYPFAF 133
 DB 72 TVGVGGAYALVRAVDELTRRNIIQQGSLRKLVIHLSGLFLVAMPFSGSTEARYPFAF 131
 QY 134 VPLVNGLRVINGLSISPSNMLIKSVTRGRABELLKGPLPYVALIFSAVFPWRBSPIG 193
 DB 132 VPLVNGLRVINGLSISPSNMLIKSVTRGRABELLKGPLPYVALIFSAVFPWRBSPIG 191
 QY 194 MISLAMCGDGIADIMGRKFGSTKIPYPRKSMAGSISMPFEGFISIALIYYSSGLY 253

Db 192 VISLMMCGDGIADIIGRRYGMKIPLYNHHKSLAGSMVLVFGVLSIGMLYYYSVLGH 251
 Qy 254 LHMWETTLQVAVMSVATVSESPIPDQDDNISVPLATITLAAYLSF 302
 Db 252 VOLDMASTLPRVAFISFVATLVESLPITKVVDNISVPLATMAVAFFTE 300

RESULT 10 US-10-634-548-46

Sequence 46, Application US/10634548
 Publication No. US20040045051A1
 GENERAL INFORMATION:
 APPLICANT: No. US20040045051A1r1s, Susan R
 APPLICANT: Lincoln, Kim
 APPLICANT: Abad, Mark Scott
 APPLICANT: Eltere, Robert
 APPLICANT: Hartsuyker, Karen Kindle
 APPLICANT: Hirschberg, Joseph
 APPLICANT: Karunanandaa, Balasulojini
 APPLICANT: Moohiri, Farhad
 APPLICANT: Stein, Joshua C.
 APPLICANT: Valentin, Henry E.
 APPLICANT: Venkatesh, Tyamagondlu V.
 TITLE OF INVENTION: Tocopherol biosynthesis related genes and used thereof
 FILE REFERENCE: Ren-01-125
 CURRENT APPLICATION NUMBER: US/10/634,548
 PRIOR FILING DATE: 2003-08-05
 PRIOR APPLICATION NUMBER: us 60/400,689
 NUMBER OF SEQ ID NOS: 79
 SOFTWARE: PatentIn version 3.2
 SEQ ID NO 46
 LENGTH: 314
 TYPE: PRT
 ORGANISM: Oryza sativa
 US-10-634-548-46

Query Match 51.5%; Score 791.5; DB 4; Length 314;
 Best Local Similarity 51.1%; Pred. No. 8.2e-74;

Matches 159; Conservative 60; Mismatches 81; Indels 11; Gaps 4;

Qy 2 AATPLSPINHQLR---FGNNSL-----TTRFCSPGFLISSPCFIGHT-GMGATQDRA 53
 Db 3 AAAPVDVVRHFPCCSSVVAASSSLLSRKSLSPAAAASMSRRRLVGVGAAA---A 59
 Qy 54 RRLISSAVATNSLIHDVGATVAVLGAYALVLSFESITKXNVIOQSISRLVHILSGLL 113
 Db 60 PAVAAALASATPALRDCATLITAGAYSLVRAFDGLTARLLIEONLSRKIVHVLSCVL 119
 Qy 114 FVLAMPFGSGTEARFPAFPLVNGRLVINGLSISNSMLISVREGRAEELKGPL 173
 Db 120 FMSSWPLFSNSTEARFPAFVPLNCRLLTYGRLSTDEALVSVREGREBELKGPL 179
 Qy 174 FVVALLFSAVFFWRRESPIGMISLMMCGDGIADIMGRKSGTKIPNPKRSWAGSISM 233
 Db 180 YVIVLVSVLVFWRQSPIGIVSLMSMGSGDPADIVGRYGSAKLPENKSKWIGSISM 239
 Qy 234 FIFGFFISIALIYYSSIGYIHMWETTLQVAVMSVATVSESPIPDQDDNISVPLA 293
 Db 240 FISGFLLSALMLFFSCIGYFTVCMDLALGKLALVALAATVVEICPVNDVVDNISVPLA 299
 Qy 294 TILAAYLSFGY 304
 Db 300 TMLAAYLLFGY 310

RESULT 11
 US-10-425-114-55785
 Sequence 55785, Application US/10425114
 Publication No. US20040034888A1
 GENERAL INFORMATION:
 APPLICANT: Liu, Jingdeng
 APPLICANT: Zhou, Yihua

APPLICANT: Kovalic, David K.
 APPLICANT: Screen, Steven E
 APPLICANT: Tabaska, Jack E
 APPLICANT: Cao, Yongwei
 TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
 FILE REFERENCE: 38-21(5313)B
 CURRENT APPLICATION NUMBER: US/10/425,114
 PRIOR FILING DATE: 2003-04-28
 NUMBER OF SEQ ID NOS: 73128
 SEQ ID NO 55785
 LENGTH: 332
 TYPE: PRT
 ORGANISM: Oryza sativa nipponbare
 FEATURE:
 OTHER INFORMATION: Clone ID: LIB3432-014-E9_FLI.pep
 US-10-425-114-55785

Query Match 51.5%; Score 791.5; DB 4; Length 332;
 Best Local Similarity 51.1%; Pred. No. 8.9e-74;
 Matches 159; Conservative 60; Mismatches 81; Indels 11; Gaps 4;

Qy 2 AATPLSPINHQLR---FGNNSL-----TTRFCSPGFLISSPCFIGHT-GMGATQDRA 53
 Db 21 AAAPVDVVRHFPCCSSVVAASSSLLSRKSLSPAAAASMSRRRLVGVGAAA---A 77
 Qy 54 RRLISSAVATNSLIHDVGATVAVLGAYALVLSFESITKXNVIOQSISRLVHILSGLL 113
 Db 78 PAVAAALASATPALRDCATLITAGAYSLVRAFDGLTARLLIEONLSRKIVHVLSCVL 137
 Qy 114 FVLAMPFGSGTEARFPAFPLVNGRLVINGLSISNSMLISVREGRAEELKGPL 173
 Db 138 FMSSWPLFSNSTEARFPAFVPLNCRLLTYGRLSTDEALVSVREGREBELKGPL 197
 Qy 174 FVVALLFSAVFFWRRESPIGMISLMMCGDGIADIMGRKSGTKIPNPKRSWAGSISM 233
 Db 198 YVIVLVSVLVFWRQSPIGIVSLMSMGSGDPADIVGRYGSAKLPENKSKWIGSISM 257
 Qy 234 FIFGFFISIALIYYSSIGYIHMWETTLQVAVMSVATVSESPIPDQDDNISVPLA 293
 Db 258 FISGFLLSALMLFFSCIGYFTVCMDLALGKLALVALAATVVEICPVNDVVDNISVPLA 317
 Qy 294 TILAAYLSFGY 304
 Db 318 TMLAAYLLFGY 328

RESULT 12 US-10-634-548-49

Sequence 49, Application US/10634548
 Publication No. US20040045051A1
 GENERAL INFORMATION:
 APPLICANT: No. US20040045051A1r1s, Susan R
 APPLICANT: Lincoln, Kim
 APPLICANT: Abad, Mark Scott
 APPLICANT: Eltere, Robert
 APPLICANT: Hartsuyker, Karen Kindle
 APPLICANT: Hirschberg, Joseph
 APPLICANT: Karunanandaa, Balasulojini
 APPLICANT: Moohiri, Farhad
 APPLICANT: Stein, Joshua C.
 APPLICANT: Valentin, Henry E.
 APPLICANT: Venkatesh, Tyamagondlu V.
 TITLE OF INVENTION: Tocopherol biosynthesis related genes and used thereof
 FILE REFERENCE: Ren-01-125
 CURRENT APPLICATION NUMBER: US/10/634,548
 PRIOR FILING DATE: 2003-08-05
 PRIOR APPLICATION NUMBER: us 60/400,689
 NUMBER OF SEQ ID NOS: 79
 SOFTWARE: PatentIn version 3.2
 SEQ ID NO 49
 LENGTH: 657

```

QY TYPE: PRT
; ORGANISM: Oryza sativa
US-10-634-548-49

Query Match          51.5%; Score 791.5; DB 4; Length 657;
Best Local Similarity 51.1%; Pred. No. 2.3e-73;
Matches 159; Conservative 60; Mismatches 81; Indels 11; Gaps 4;

OY      2 AATPLSPINHOLCR---FGNNSL-----TTHRCSPFLISSPCFGLT-GMGSAOTLRA 53
DB      346 AAARPVDVVRHPPCCSSVAASSSLLSRKSKRLASPAAMAAASMRRLVLGVGAAA--A 402

OY      54 RRSLSAVATNSLHDVGATVAVLGAAVLYLPESLTGRNVIQOSLSRKLVIILSGLL 113
DB      403 PAVALMAASATPALMDCACTTLITAGAYSLVAAPDCLTRRLLEQLSKRIYHVLGVL 462

OY      114 FYLAAMPDFSGSTEARYFAAFVPLVNGLRVLVINGLSISPNSMLIKSVTREGRAEELKGPL 173
DB      463 FMSWNPFLFNSSTEARFPFAIVPLPNCIRLLTYGLRISTDEALVKSVTREGEPEELLRGPL 522

OY      174 FYVALLESADVFWRRSPICGMISLAMCGGDGIADIMGRKPGSKTIYYNRPKXNASGISM 233
DB      523 YVIVLVTVLVSUVMWRSPIGIVSLSMWSGGDGFDIIVGREYGSXKLPFNKRKSWIGISM 582

OY      234 FTGEGFTISALLYYYSLSGLIGHNMWETTGORVAMVSVAIVTESLPTTDOLDNDNISVPLA 293
DB      583 FTSGFLSLAMLFFFSCLGFETVCMDLAKLVALALTAATVECI PVNDVDDNISVPLA 642

OY      294 TILAAYLSFGY 304
DB      643 TWLAAYLTFGY 653

RESULT 13
US-10-437-963-179783
; Sequence 179783, Application US/10437963
; Publication No. US2004012343A1

GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
APPLICANT: Wu, Wei
APPLICANT: Boukharov, Andrey A.
APPLICANT: Barbazuk, Brad
APPLICANT: Li, Ping
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
FILE REFERENCE: 38-21(53221)B
CURRENT APPLICATION NUMBER: US/10/437, 963
NUMBER OF SEQ ID NOS: 204966
SEQ ID NO 179783
LENGTH: 657
TYPE: PRT
ORGANISM: Oryza sativa
FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_77211C.1.pep
US-10-437-963-179783

Query Match          51.5%; Score 791.5; DB 4; Length 657;
Best Local Similarity 51.1%; Pred. No. 2.3e-73;
Matches 159; Conservative 60; Mismatches 81; Indels 11; Gaps 4;

OY      2 AATPLSPINHOLCR---FGNNSL-----TTHRCSPGFLISSPCFGLT-GMGSAOTLRA 53
DB      346 AAARPVDVVRHPPCCSSVAASSSLLSRKSKRLASPAAMAAASMRRLVLGVGAAA--A 402

OY      54 RRSLSAVATNSLHDVGATVAVLGAAVLYLPESLTGRNVIQOSLSRKLVIILSGLL 113
DB      403 PAVALMAASATPALMDCACTTLITAGAYSLVAAPDCLTRRLLEQLSKRIYHVLGVL 462

OY      114 FYLAAMPDFSGSTEARKYFAAFVPLVNGLRVLVINGLSISPNSMLIKSVTREGABEELKKPL 173

```

QY	DB	Query Match	Best Local Similarity	Matches	51.5%; Score 791.5; DB 4; Length 803; Pred. No. 3e-73; Mismatches 81; Indels 11; Gaps 4;
QY	Db	174 FYVALLFSAVFWRPESPIGMSILAMCGDGIADIGRKGFSYKIDYNPKKSWAGSISM 233	463 FMSSWPFNSSTERFPAALVPLINCRLITTYGRLSTDEALVKSVTRREGKEBELLGPL 522		
QY	Db	523 YVIVLVSVLVFMRQSPGIVLSIMWSGGGFADIVGRRGSKLCPFNEKSWIGSISM 582			
QY	Db	234 FIFGFFISIALLYYSSSLGYLHNMWETTLQRYAVMSVAVTVYSELPTTDOLDNISVPLA 293			
QY	Db	583 FIFGFFISIALMLFFSCIGYFTVCMDLALGKLALVALAIVVECTIPVNDVVDNISVPLA 642			
QY	Db	294 TILAAYLSFGY 304			
QY	Db	643 TMLAAYLLFGY 653			
QY	Db	294 TILAAYLSFGY 304			
QY	Db	729 FIFGFFISIALMLFFSCIGYFTVCMDLALGKLALVALAIVVECTIPVNDVVDNISVPLA 788			
QY	Db	294 TILAAYLSFGY 304			
QY	Db	789 TMLAAYLLFGY 799			

RESULT 15

US-10-437-963-179784

Sequence 179784, Application US/10437963

Publication No. US20040123343A1

GENERAL INFORMATION:

APPLICANT: LA ROBA, THOMAS J.

APPLICANT: Kovalic, David K.

APPLICANT: Zhou, Yihua

APPLICANT: Cao, Yongwei

APPLICANT: Wu, Wei

APPLICANT: Boukharov, Andrey A.

APPLICANT: Barbazuk, Brad

APPLICANT: BALDARUK, ID Pina

TITLE OR INVENTION: Polye Nucleic Acid Molecules and Other Molecules Associated With

TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules

PLANTS AND USES THEREOF

FILE REFERENCE: 38-21(53221)B

CURRENT APPLICATION NUMBER:

CURRENT FILING DATE: 2003-03-20

; NUMBER OF SE

; SEQ ID NO 179

; LENGTH: 803

TYPE: PRT

ORGANISM: *Oryza sativa*

FEATURE:

OTHER INFORMATION: C10ne ID: PAT_MRT4530_77212C.1.pep
US-10-437-963-179784

Query Match	51.5%	Score 791.5;	DB 4;	Length 803;
-------------	-------	--------------	-------	-------------

Matches 159; Conservative 60; Mismatches 81; Indels 11; Gaps 4;

2 AATLPLSPINHQLCR---FGNSL---TTHRFCSPGFLISSPCFIGLT-GMGSATQLRA 53

Db 492 A A A R P V D V R H F P C S S S V A A S S S L L S R S K S R I A S P A A A A A S M R R R L V L G V G A A - - - A 546

54 RRSLTSSAVATNSI.LHDVGATVAVI.GGAYAI.VT.SPEST.TKRNVIQOSI.SPKI.VHTI.SGI.T. 113

E40 PIVAT AASTUVAI BDCAAHT I TEICVETIMABPCIT EMBBY TEOUF CNYTTHN COVF CAOS
::|||:|||::|||::|||::|||::|||:
DH

[illegible][illegible]

QY 174 FYVLALFSAVPFWRESPIGMISLAMCGGDGIADIMGRKFGSTKI PYNPRKSWAGSISM 233

Db 669 YVIVLVSVLFWRQSPIGIVLSMSMGDFADIVGRRYGSAKL.PFNENKSWIGISM 726

234 RIFGFFISIALYYSSLSGLHMNETTLORVAMVSMVATVESLPITDOLDNISVPLA 293

729 PTSGPTLSAIMEVPSCTGYETVUCWDIALGKIALVALAATVECTPVNDVDNDNISVPIA 786

3: 504 ETT 3307 504

Search completed: March 16, 2006, 19:53:26
Job time : 80 secs

Job time : 80 secs

THIS PAGE LEFT BLANK

US-11-096-568A-19560
 ; Sequence 19560, Application US/11096568A
 ; Publication No. US20060048240A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Alexandrov, Nikolai et al.
 ; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
 ; FILE REFERENCE: 2750-1592PUS2
 ; CURRENT APPLICATION NUMBER: US/11/096,568A
 ; CURRENT FILING DATE: 2005-04-01
 ; NUMBER OF SEQ ID NOS: 34471
 ; SEQ ID NO 19560
 ; LENGTH: 268
 ; TYPE: PRT
 ; ORGANISM: Zea mays subsp. mays
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (1)..(268)
 ; OTHER INFORMATION: Ceres Seq. ID no. 12373618
 US-11-096-568A-19560

Query Match 50.5%; Score 776; DB 7; Length 268;
 Best Local Similarity 58.8%; Pred. No. 9,4e-62;
 Matches 143; Conservative 52; Mismatches 48; Indels 0; Gaps 0;

QY 60 SAVATNSLHDVGATVAVLGGAYALVLSFESLTKRNVIOQSLSRKLVHLSGLFVLAMP 119
 DB 18 AAAAPPAVLQDGAATVLLTAGVSLVRVDELTERLLEKLSRKVVHLSGLFVMSWP 77
 QY 120 IFSGSTEARYFAAFVPLVNGRLVINGLSISPSNMLIKSVTREGRAEELKGPLFYVAL 179
 DB 78 LFSNSTEARFYFAAVVPLFINSKMLLYGLRLTYDBALVKSVTREGKEELRGPLYVVAL 137
 QY 160 LFSAPFPRFESTIGMISLAMCGGDIADIMGRKGSITKIPNPKKSWAGSISMTIFPF 239
 DB 138 LFSVAVFWRPESIGIVLSMMSGDGFADIVGRYGSALPFRKKSAGSISMTISGFL 197
 QY 240 ISALLYYYSLSGLYHMMWETTLQRVAMSVAVTVESLPTDQDDNISVPLATILAAV 299
 DB 198 LSAMMMLYFSSLGIVDIWEALGKLVLAALVAVVECPVTEVVDNISVPLATMLVAF 257

QY 300 LSF 302
 DB 258 LLF 260

RESULT 3
 US-11-096-568A-19561
 ; Sequence 19561, Application US/11096568A
 ; Publication No. US20060048240A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Alexandrov, Nikolai et al.
 ; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
 ; FILE REFERENCE: 2750-1592PUS2
 ; CURRENT APPLICATION NUMBER: US/11/096,568A
 ; CURRENT FILING DATE: 2005-04-01
 ; NUMBER OF SEQ ID NOS: 34471
 ; SEQ ID NO 19561
 ; LENGTH: 196
 ; TYPE: PRT
 ; ORGANISM: Zea mays subsp. mays
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (1)..(196)
 ; OTHER INFORMATION: Ceres Seq. ID no. 12373619
 US-11-096-568A-19561

Query Match 41.4%; Score 636; DB 7; Length 196;
 Best Local Similarity 61.8%; Pred. No. 1.8e-49;
 Matches 115; Conservative 38; Mismatches 33; Indels 0; Gaps 0;
 QY 117 AMPIFSSTEARYFAAFVPLVNGRLVINGLSISPSNMLIKSVTREGRAEELKGPLFYV 176

DB :||| ||||| ||| :||| :||| ||||| :||| |||||
 3 SMPFSSNSTEARFYFAAVVPLFINSKMLLYGLRLTYDBALVKSVTREGKEELRGPLYV 62
 QY 177 LALFSAVPFWRPESIGMISLAMCGGDIADIMGRKGSITKIPNPKKSWAGSISMTIF 236
 DB 63 LVLFSAVPFWRPESIGIVLSMMSGDGFADIVGRYGSALPFRKKSAGSISMTIS 122
 QY 237 GPFISALLYYYSLSGLYHMMWETTLQRVAMSVAVTVESLPTDQDDNISVPLATIL 296
 DB 123 GFLLSAMMMLYFSSLGIVDIWEALGKLVLAALVAVVECPVTEVVDNISVPLATML 182
 QY 297 AAYLSF 302
 DB 183 VAFLLF 188

RESULT 4
 US-11-096-568A-1771
 ; Sequence 1771, Application US/11096568A
 ; Publication No. US20060048240A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Alexandrov, Nikolai et al.
 ; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
 ; FILE REFERENCE: 2750-1592PUS2
 ; CURRENT APPLICATION NUMBER: US/11/096,568A
 ; CURRENT FILING DATE: 2005-04-01
 ; NUMBER OF SEQ ID NOS: 34471
 ; SEQ ID NO 1771
 ; LENGTH: 259
 ; TYPE: PRT
 ; ORGANISM: Zea mays subsp. mays
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (1)..(259)
 ; OTHER INFORMATION: Ceres Seq. ID no. 15178975
 US-11-096-568A-1771

Query Match 39.9%; Score 613; DB 7; Length 259;
 Best Local Similarity 43.9%; Pred. No. 2.8e-47;
 Matches 125; Conservative 50; Mismatches 54; Indels 56; Gaps 4;

QY 29 SPGFLL--SSPCFGLTGMGSATQLRARSLIS-----SAVATNSLHDVGATVAV 77
 DB 12 SPNSLLSRSPPHAVALPSPGSSWR--RRLLVGTVPAAVLAALAAAPPAVLQDGAATVLI 70
 QY 78 LGAYALVLSFESLTKRNVIOQSLSRKLVHLSGLFVLAMPFSGSTEARYFAAFVPLV 137
 DB 71 TAGVSLVRVDELTERLLEKLSRKVVHLSGLFVMSWPLF----- 114
 QY 138 NGLRLVINGLSISPSNMLIKSVTREGRAEELKGPLFYVALLFSAPFWRPESIGIVSL 197
 DB 115 -----RELLRGPLYVVALFVLFWRPESIGIVSL 146
 QY 198 AMMCGGDIADIMGRKGSITKIPNPKKSWAGSISMTIFGPFISALLYYYSLSGLYHMM 257
 DB 147 SMMSGDGFADIVGRYGSALPFRKKSAGSISMTISGFLLSAMMMLYFSSLGIVDI 206
 QY 258 WETTLQRVAMSVAVTVESLPTDQDDNISVPLATILAAVLSF 302
 DB 207 WEALGKLVLAALVAVVECPVTEVVDNISVPLATMLVAFLLF 251

RESULT 5
 US-11-096-568A-1770
 ; Sequence 1770, Application US/11096568A
 ; Publication No. US20060048240A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Alexandrov, Nikolai et al.
 ; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
 ; FILE REFERENCE: 2750-1592PUS2
 ; CURRENT APPLICATION NUMBER: US/11/096,568A

CURRENT FILING DATE: 2005-04-01
NUMBER OF SEQ ID NOS: 34471
SEQ ID NO 1770
LENGTH: 270
TYPE: PRT
ORGANISM: Zea mays subsp. mays
FEATURE:
NAME/KEY: misc feature
LOCATION: (1)-(270)
OTHER INFORMATION: Cereals Seq. ID no. 15178974
US-11-096-568A-1770

Query Match 39.9%; Score 613; DB 7; Length 270;
Best Local Similarity 43.9%; Pred. No. 3e-47;
Matches 125; Conservative 50; Mismatches 54; Indels 56; Gaps 4;

QY 29 SPFGFLI--SSPCFGLTGMGSAATOLRARSLIS-----SAVATNSLHDVATVAV 77
DB 23 SPNSLLSRSPHAAVLAPSPGSSMR--RRLILGVGTVAALAAAAPAVLQDGAVTYLI 81
QY 78 LGGAYALVSPESLTKRNVIOQSLRKLVHLSGLLFTLAMPISGSTEAFYFAFVPLV 137
DB 82 TAGAYSLVRVDELTERRLIKSLRKVHVLSGLVFMSSWPLF----- 125
QY 138 NGRLVINGLSISPSNMLIKSVTRGRAEELKGPLFYVALLFSAVFWMRESPIGMI 197
DB 126 -----RELKGPLYVALVLSVLFVWMRESPIGIVSL 157
QY 198 AMMGCGDGIADIMGRKFGSTIKIPYPRKSWAGSISMPFEGFISIALLYYSSGLYLMN 257
DB 158 SMSGCGDGFADIVGRRRYSALLPFRKSKWSAGSISMPISGFLLSMMMLYSSGLYDVI 217
QY 258 WETLQVAMVSMVATVYESLPITDLDNISVPLATILAYLSF 302
DB 218 WEELGKLALVALAAVVECVPTVEVDNISVPLATVAFLLF 262

RESULT 6

US-11-096-568A-1772
Sequence 1772, Application US/11096568A
Publication No. US20060048240A1
GENERAL INFORMATION:
APPLICANT: Alexandrov, Nikolai et al.
TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
TITLE OF INVENTION: Theby
FILE REFERENCE: 2750-1592PUS2
CURRENT APPLICATION NUMBER: US/11/096,568A
CURRENT FILING DATE: 2005-04-01
NUMBER OF SEQ ID NOS: 34471
SEQ ID NO 1772
LENGTH: 224
TYPE: PRT
ORGANISM: Zea mays subsp. mays
FEATURE:
NAME/KEY: misc feature
LOCATION: (1)-(224)
OTHER INFORMATION: Cereals Seq. ID no. 15178976
US-11-096-568A-1772

Query Match 39.7%; Score 610; DB 7; Length 224;
Best Local Similarity 47.7%; Pred. No. 4.4e-47;
Matches 116; Conservative 45; Mismatches 38; Indels 44; Gaps 1;

QY 60 SAVATNSLHDVATVAVLGGAYALVSPESLTKRNVIOQSLRKLVHLSGLLFTLAMP 119
DB 18 AAAAPPAVLQDGAVTYLIAGAYSLVRVDELTERRLIKSLRKVHVLSGLVFMSSWP 77
QY 120 IFSGSTEARVFAAFVPLVNGRLVINGLSISPSNMLIKSVTRGRAEELKGPLFYVAL 179
DB 78 LF-----RELKGPLYVALV 93
QY 180 LFSAVFWMRESPIGMISSLAMCGDGIADIMGRKFGSTIKIPYPRKSWAGSISMPFEGF 239

DB 94 LFSVLFWMRESPIGIVLSMWSGGDGFADIVGRRRYSALKLPFRKSKWSAGSISMPISGFL 153
QY 240 ISIALIYSSGLGLHANNWETTLQVAMVSMVATVYESLPITDLDNISVPLATILAY 299
DB 154 LSAAMMYFSSLYDVIWEELGKLALVALAAVVECVPTVEVDNISVPLATVAF 213
QY 300 LSF 302
DB 214 LSF 216

RESULT 7

US-11-096-568A-14812
Sequence 14812, Application US/11096568A
Publication No. US20060048240A1
GENERAL INFORMATION:
APPLICANT: Alexandrov, Nikolai et al.
TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
TITLE OF INVENTION: Theby
FILE REFERENCE: 2750-1592PUS2
CURRENT APPLICATION NUMBER: US/11/096,568A
CURRENT FILING DATE: 2005-04-01
NUMBER OF SEQ ID NOS: 34471
SEQ ID NO 14812
LENGTH: 302
TYPE: PRT
ORGANISM: Zea mays subsp. mays
FEATURE:
NAME/KEY: misc feature
LOCATION: (1)-(302)
OTHER INFORMATION: Cereals Seq. ID no. 12336654
US-11-096-568A-14812

Query Match 36.6%; Score 562.5; DB 7; Length 302;
Best Local Similarity 41.4%; Pred. No. 1.1e-42;
Matches 115; Conservative 53; Mismatches 101; Indels 9; Gaps 4;

QY 28 CSPGFLISSPCFGLTGMGSAATOLRARSLISSAV---ATNSLHDVATVAVLGGAYAL 84
DB 31 CAP---RSLCFRRRRSSRLAAE--RTRPPTMAAISLEAGALAHDDLSAVVTGGVALAL 85
QY 85 VLSFESLTKRNVIOQSLRKLVHLSGLLFTLAMPISGSTEAFYFAFVPLVNGRLVYI 144
DB 86 LKPFEBELAKGVFPQKSLRKLVHLSGLVFLFPLBSSGYTFPLAALAGVNIIRLL 145
QY 145 NGLSISPSNMLIKSVTRGRAEELKGPLFYVALLFSAVFWMRESPIGMISSLAMCGD 204
DB 146 LGLGLMKNEMAVKMSRSGDYRELLKGPLYAAATITPATISLMRTSPVALIALICNLGAGD 205
QY 205 GIADIMGRKFGSTIKIPYPRKSWAGSISMPFEGFISIALLYYSSGLYLMNWTTLQ 264
DB 206 GIADIVGRRLKELPYNPNKSYAGSIAMAVAGFLASGVNHYRTGFIETWYALS- 264
QY 265 VAMVSMVATVYESLPITDLDNISVPLATILAYLSF 302
DB 265 PLVVSVAALVESHPISTELDDNITVLTLSPLVGSLLF 302

RESULT 8

US-11-096-568A-14813
Sequence 14813, Application US/11096568A
Publication No. US20060048240A1
GENERAL INFORMATION:
APPLICANT: Alexandrov, Nikolai et al.
TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
TITLE OF INVENTION: Theby
FILE REFERENCE: 2750-1592PUS2
CURRENT APPLICATION NUMBER: US/11/096,568A
CURRENT FILING DATE: 2005-04-01
NUMBER OF SEQ ID NOS: 34471
SEQ ID NO 14813
LENGTH: 247
TYPE: PRT

Db 1 MMSGDDGADIVGRRRYSAGKLPFNRRKSWAGSISMFISGFLLSAMMLYFSSLGIVDTW 60
Qy 259 ETTQRYAMVSMVATVSESLPTDQDNDNISVPLATITLAATYSP 302
Db 61 EALGKALVALVATVCEVTEVVDNISVPLATMLVAFILF 104

RESULT 12
US-11-096-568A-14320
; Sequence 14320, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nickolai et al.
; TITLE OF INVENTION: Sequence-determined DNA fragments and corresponding polypeptides
; FILE REFERENCE: 2/50-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 14320
; LENGTH: 111
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)_(111)
; OTHER INFORMATION: Ceres Seq. ID no. 3626197
US-11-096-568A-14320

Query Match 22.2%; Score 341; DB 7; Length 111;
Best Local Similarity 59.2%; Pred. No. 1,66-23;
Matches 61; Conservative 23; Mismatches 19; Indels 0; Gaps 0;

Qy 200 MCGDGDGADIVGRRRYSAGKLPFNRRKSWAGSISMFISGFLLSAMMLYFSSLGIVDTW 259
Db 1 MMSGDDGADIVGRRRYSAGKLPFNRRKSWAGSISMFISGFLLSAMMLYFSSLGIVDTW 60

Qy 260 TTQRYAMVSMVATVSESLPTDQDNDNISVPLATITLAATYSP 302
Db 61 EALGKALVALVATVCEVTEVVDNISVPLATMLVAFILF 103

RESULT 13
US-11-087-099-1186
; Sequence 1186, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53450)B RP
; CURRENT APPLICATION NUMBER: US/11/087,099
; CURRENT FILING DATE: 2005-03-22
; NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 1186
; LENGTH: 477
; TYPE: PRT
; ORGANISM: Chlorobium tepidum T1S
US-11-087-099-1186

Query Match 8.6%; Score 132.5; DB 7; Length 477;
Best Local Similarity 28.0%; Pred. No. 0,00033;
Matches 53; Conservative 37; Mismatches 74; Indels 25; Gaps 11;

Qy 129 YFAA-PVPLVNGRLVINGLSISPNMLIKS-----VTRGRABELIK-GPLFYVAL 180
Db 11 YFSNFYPLAALFPAVVG-LVSLAGILQLHGPVVTGG--ERVISGYPVLFPVLF 67

Qy 181 FSAVFPWESP-IGMISLMMCGDGDIVGRRRYSAGKLPFNRRKSWAGSISMFISGFLLSAMMLYFSSLGIVDTW 259
Db 68 LQALFLMEHWIQLISMLVIGIDALALVGTAAAGSHIENLTYSKISGSMAMPIS 127

Qy 238 FFI-SIALLYYSSL-GYLHNMETTLQRYAMVSMVATVSESLPTDQDNDNISVPLATITLAATYSP 302
Db 61 EALGKALVALVATVCEVTEVVDNISVPLATMLVAFILF 103

Db 128 LVIVSIFVRDPAFTGVLGQPIWKLLALALLAVTAVAL-----LSMGIDNLP 182
Qy 292 LATILAAT 300
Db 183 LAIAYVLYV 191

RESULT 14
US-10-793-626-2168
; Sequence 2168, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STRAPHYLOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: P03480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 2168
; LENGTH: 383
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
US-10-793-626-2168

Query Match 6.4%; Score 98.5; DB 6; Length 383;
Best Local Similarity 19.9%; Pred. No. 0,27; Indels 117; Gaps 13;
Matches 67; Conservative 49; Mismatches 103; Indels 117; Gaps 13;

Qy 13 QLCFRGNSLTTHRCFSPGFLISSPCFGLTGMSATQARRSLISSAVATNSLHDVG 72
Db 61 QIVFESKSTLNLGIFET---IISNP---WLVGLGQFILLITVIMLGVEKG--LEKAS 112

Qy 73 ATYAVLGCAVLYVSFESLTGRVNIQ-----QSLSKLVHISGLFLVLAAPTES 122
Db 113 KIMPLLEFLIIVVAOSTLLEGALBGRVYLLQPRVEDMS-----IQGLFALGQSFET 166

Qy 123 ---GSTEARVPAFV-----LVNGRLVINGLSISPNMLIKS-----VTRGRABELIK-GPLFYVAL 180
Db 167 LSLGTTGMITYASAPKMTIKSALSIYVNNILISVLAGLAIFALKTBTGQEG--- 223

Qy 167 ELKGPFLYVALLPFAVFPWESP-IGMISLMMCGDGDIVGRRRYSAGKLPFNRRKSWAGSISMFISGFLLSAMMLYFSSLGIVDTW 259
Db 224 ---PGLLPKVLPLVSEWTF----- 240

Qy 227 WAGSISMFISGFLLSAMMLYFSSLGIVDTW 259
Db 241 --GTFYFIFLPLFAL--TSSISLLELVNSFTKNDNSKRVKVALIGSLVFIISIP 296

Qy 280 IT-----DQDNDNISVPLATITLAATYSP 302
Db 297 ATLSFSSLSHLRFAGATTFDMDFTVSNILPLGAL 332

RESULT 15
US-11-092-140-6
; Sequence 6, Application US/11092140
; Publication No. US20050262590A1
; GENERAL INFORMATION:
; APPLICANT: Subramaniam, S.; Slater, S.; Karberg, K.; Chen, R.; Valentini, H.; Wong, Y.
; TITLE OF INVENTION: Nucleic Acid Sequences to Proteins Involved in Tocopherol Synthesis
; FILE REFERENCE: 16515,054
; CURRENT APPLICATION NUMBER: US/11/092,140
; CURRENT FILING DATE: 2005-03-29
; PRIOR APPLICATION NUMBER: US/09/688,069
; PRIOR FILING DATE: 2000-10-14
; NUMBER OF SEQ ID NOS: 114
; SEQ ID NO 6

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - nucleic search, using frame_p2n model

Run on: March 16, 2006, 19:54:00, Search time 7741 Seconds
(without alignments)
2232.323 Million cell updates/sec

Title: US-10-634-548-2

Perfect score: 1536
Sequence: 1 MAATLPLSPINHQCRCRGN.....DDNISVPLATLIAAYLSFGY 304

Scoring table:

BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xjh
-Q=/abs/ABSSMBR.spool/US10634548/runat_16032006_084432_10307/app_query.fasta-1
-DB=Genemb1 -QFMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-LOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=apct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000 -HOST=abs02h
-USER=US10634548 @CGN_1_14939_@runat_16032006_084432_10307 -NCPU=6 -ICPU=3
-NO_MMAP -NEG_SCORES=0 -WAIT -DSPBLOC=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Genemb1.*

1: gb_ba.*
2: gb_in.*
3: gb_env.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pr.*
9: gb_ro.*
10: gb_atg.*
11: gb_ey.*
12: gb_un.*
13: gb_vl.*
14: gb_hcg.*
15: gb_pl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1536	100.0	1004	15 BT021123	Arabidops
2	1531	99.7	1072	15 BT004006	Arabidops
3	1523	99.2	1091	15 AY085036	Arabidops

RESULT 1	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	FEATURES
BT021123	Arabidops thaliana	At5g04490 gene, complete cde.	BT021123	BT021123.1	GI:60543320	Arabidops thaliana (thale cress)	Kim,C.J., Chen,H., Cheuk,R., Shinn,P. and Ecker,J.R.	Submitted (05-MAR-2005)	Salk Institute Genomic Analysis Laboratory (SIGNAL), Plant Biology Laboratory, The Salk Institute for Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA	Location/Qualifiers	
1	1004	bp	linear	PLN 05-MAR-2005							1..1004
											/organism="Arabidops thaliana"
											/mol_type="mRNA"

submitted to Genbank.

FEATURES
Location/Qualifiers
1..1072
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosome="5"
/clone="RAP15-38-N13 (R50070)"
/ecotype="Columbia"
/note="This clone is in a modified pBluescript vector2 (lambda ps) as a BamHI/XhoI insert."
1..1072
/gene="At5g04490"
1..69
/gene="At5g04490"
70..984
/gene="At5g04490"
/codon_start=1
/evidence=experimental
/product="unknown protein"
/protein_id="AA042044.1"
/db_xref="GI:28393229"
/translation="MAATLPSPINHOLCRFNNSLTTHRFCSPPFLISSPCFGLTG
MGSATQLRARRSLISAATNSLDHGAVALVAGAYALVLSPESTLKNVQOOS
RKVHLISGLFLVLPAPIPSGSTEARPAAPVPLVNLRLVINGLSISPMILKSYT
RGRARERLKGPLPYVALPSAIPFRESPIGMSIAMMGGDGDIADIMGRKGSYK
IPNPKRSNMGSIEMTFGFFISIALYYSISGLYLMNMTTLQKRVAMVMAVAVE
SLPTDQLDNISVPLATILAAVLSFGY"
misc_difference
289
/gene="At5g04490"
/note="compared to genomic sequence resulting in an amino
acid sequence difference"
/replace="a"
985..1072
/gene="At5g04490"

ORIGIN
3' UTR

Alignment Scores:
Pred. No.: 2.6e-128 Length: 1072
Score: 1531.00 Matches: 303
Percent Similarity: 99.7% Conservative: 0
Best Local Similarity: 99.7% Mismatches: 1
Query Match: 99.7% Indels: 0
DB: 15 Gaps: 0

US-10-634-548-2 (1-304) x BT004006 (1-1072)

Qy 1 MetAlaAlaThrLeuProLeuSerProIleAsnIleGlnLeuCyArpPheGlyAsnAsn 20
Db 70 ATGGCAGCACTTAACCTCTATCTCCATCAATCAATCGTTGTCCGTCGGGAACAC 129
Qy 21 SerLeuThrThrAlaArgPheCySerProGlyPheLeuIleSerSerProCyAspIle 40
Db 130 TCTTTGACGACTCAACCGCTCTGTTCTCCGCTTCTTGATTTCTTCCCTGTTTCA 189
Qy 41 GlyLeuThrGlyMetGlySerAlaThrGlnLeuArgAlaArgArgSerLeuIleSerSer 60
Db 190 GGTTCACCGAAGGGCTGCTACACAGTTAGTGCTGCTGCTGCTGCTGCTGCTTCA 249
Qy 61 AlaValAlaThrAsnSerLeuLeuAspValGlyAlaThrValAlaValLeuGlyGly 80
Db 250 GCAGTTCGACGAATTCGCTTTCATGACGTCGAGCCGCGTGGCAGTCTTGATGGA 309
Qy 81 AlaThrAlaLeuValLeuSerPheGluSerLeuThrIlyAspAsnValIleGlnGlnSer 100
Db 310 GCATTAACCGCTTGTCTTAAGCTTCGAGAGTCTCAACCAAGCAATTCATCAACAGT 369
Qy 101 LeuSerArgIysLeuValHisIleLeuSerGlyLeuLeuPheValLeuAlaTrpProIle 120
Db 370 TTGAGCGAAGAGCTTGTCATATCTCTCAGGTCTGCTTTCGTACTTGCGTGCACATC 429
Qy 121 PheSerGlySerThrGlnAlaArgIyrPheAlaIlePheValProLeuValAsnGlyLeu 140

Db 430 TTCAGCGAATGACCGAGGCTCCATCTTGCTGCTTTTGTTCGCTAGTGAATGCTTA 489
Qy 141 ArgLeuValIleAsnGlyLeuSerIleSerProAsnSerMetLeuIleIleSerValThr 160
Db 490 AGGCTTGTATTATTAACGACATCAATCCATTTCCCAATTCGATGCAATCAATCCGTACA 549
Qy 161 ArgGlnGlyArgAlaGlnGlnGlnLeuLeuGlyGlyProLeuPheThrValIleAlaLeuLeu 180
Db 550 AGAAGAGGAGAGCAGAGAGAGTTCCTTAAGAGTCTTGTTCGTTCAAGTTCCTTCTT 609
Qy 181 PheSerAlaValPhePheTrpArgGluSerProIleGlyMetIleSerLeuAlaMetMet 200
Db 610 TTCTCTCGGTTTCTTCTGAGAGAGTCTCCATGCGTATGATCTGTTAGCAATCATG 669
Qy 201 CysGlyGlyAspGlyIleAlaAspIleMetGlyArgIysPheGlySerThrIysIlePro 220
Db 670 TGTGTGCGCATGGAATGAGCTGATATATGAGAGTAAAGTTGGGTCAACATCACTACT 729
Qy 221 TyrAsnProArgIysSerTrpAlaGlySerIleSerMetPheIlePheGlyPhePheIle 240
Db 730 TCAACCCAGAGAGAGAGTGGGCGAGAGCAATCTTCATGTTCACTTCGCTTTTCAATC 789
Qy 241 SerIleAlaLeuLeuIlyrIlyrIySerSerLeuGlyIlyrLeuHisMetAsnTrpGlyThr 260
Db 790 TCATGCGATTAATCTTACTTACTTCAAGCTTGCGTACCTTCACATGAACCTGGAAACG 849
Qy 261 ThrLeuGlnArgValAlaMetValSerMetValAlaThrValValGluSerLeuProIle 280
Db 850 ACCTTGACAGAGAGAGCAATGAGTCAATGTCGCCACGATGAGTGCATCAACCATC 909
Qy 281 ThrAspGlnLeuAspAspAsnIleSerValProLeuAlaThrIleAlaAlaIlyrLeu 300
Db 910 ACCGATCAATTAAGCAGCAATATTTTCGTTCTGCTGCTACTATTTAGCTTATTTA 969
Qy 301 SerPheGlyTyr 304
Db 970 AGTTTCGATAT 981

RESULT 3
AY085036
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Arabidopsis thaliana (thale cress)
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 1091)
Haas,B.U., Volfovsky,N., Town,C.D., Troupkhan,M., Alexandrov,N.,
Feldmann,K.A., Flavell,R.B., White,O. and Salzberg,S.L.
Full-length messenger RNA sequences greatly improve genome
annotation
Genome Biol. 3 (6), RESEARCH0029 (2002)
12093376
2 (bases 1 to 1091)
Brover,V., Troupkhan,M., Alexandrov,N., Lu,Y.-P., Flavell,R. and
Feldmann,K.
Full-length cDNA from Arabidopsis thaliana
Unpublished
3 (bases 1 to 1091)
Brover,V., Troupkhan,M., Alexandrov,N., Lu,Y.-P., Flavell,R. and
Feldmann,K.
Direct Submission
Submitted (11-MAR-2002) Ceres, Inc., 3007 Malibu Canyon Road,
Malibu, CA 90265, USA
This clone sequence is one of 5,000 Ceres full-length cDNAs made
available to TIGR and Genbank. The following quality assessment of
this set was done by comparison with known proteins: two percent
of the clones are estimated to be 5'-truncated; less than one percent
are 3'-truncated; approximately two percent represent alternative

TITLE
JOURNAL
AUTHORS
REFERENCE
COMMENT

splice variants, including unspliced introns and spliced exons; one percent may contain premature stop codons; five percent may have frame shifts in a coding region. A sequence is considered to be 5'-truncated if it lacks the translation initiation start (ATG). A sequence is considered to be 3'-truncated if it lacks the C-terminal end of the encoded protein. Please note that these cDNA sequences are derived from the We or Lier ecotypes and therefore may contain polymorphisms when compared to sequences from Col-0. Genset carried out the library production and sequencing of the full-length clones. Ceres, Inc. carried out the clustering of the 5' sequences, selection of clones, and sequence assembly.

FEATURES

SOURCE

1..1091

/organism="Arabidopsis thaliana"

/mol_type="mRNA"

/db_xref="taxon:3702"

/clone="125255"

108..1022

/codon_start=1

/product="unknown"

/protein_id="AA061593.1"

/db_xref="GI:21537252"

/translation="MAATLPSPINHQCRGNNSLTTHRCSPGRLISSPCFGLTG

MGSATQLPARSLISSAVATNSLADGATVAVGAVALVLSFSLTKRWLQDSIS

RKLPHLISGLFLVLPWIFSGSTEARFAAFVPLVNLGLVNLGISPSMILKSYT

REGRAEELKGPFLVLPALFSAVFWRBESPIGMIISLMMCGGIDIMGRKGSYK

IPYPRKSMGSIEMFIFGPISIALYLYSSSLGYLHINNETLQRAVIVSMATVVE

SLPTDLDNDVSVPLATITLAVISFGY"

ORIGIN

Alignment Scores:

Pred. No.:	1..4e-127	Length:	1091
Score:	1523.00	Matches:	300
Percent Similarity:	99.7%	Conservative:	3
Best Local Similarity:	98.7%	Mismatches:	1
Query Match:	99.2%	Indels:	0
	15	Gaps:	0

US-10-634-548-2 (1-304) x AY085036 (1-1091)

QY 1 MetLAAATHTTProLeuSerProLLeuNHISGLInLeuCYAArgPheGlyAsnAsn 20
 DB 108 ATGGACAGAACCTTACCTCTATCTCCGATCATCATCGTGTGCTGGGAAACAC 167
 QY 21 SerLeuThrThrHLeuArgPheCySerProGlyPheLeuLeuSerSerProCyPheLe 40
 DB 168 TCTTTGACGACTCACCGGTTCTGTTCTCTGCTTCTTGAATTTCTTCTCTGTTTCA 227
 QY 41 GlyLeuThrGlyMetGlySerAlaThrGlnLeuArgAlaArgSerLeuLeuSerSer 60
 DB 228 GGTTCGACCGGAATGGGCTCTGCTACTACGTTACGTCGCTCGTCTCTGATCTCTTCA 287
 QY 61 AlAValAlaThrAnSerLeuLeuHISAPValGlyAlaThrValAlaValLeuGlyGly 80
 DB 288 GCAGTTGGAGGAATTCCTCTGTGATGATGACGTCGAGCACCGTGGAGTCTTGATGGA 347
 QY 81 AlATyAlaLeuValLeuSerPheGlySerLeuThrLySArgAsnValIleGlnGlnSer 100
 DB 348 GCATACGCGCTTGTCTTAAGCTTCGAGAGTCTCACCAAGCAAGCTATTCACAGAGT 407
 QY 101 LeuSerArgLySLeuValHISleuSerGlyLeuLeuPheValLeuLeuTrpProIle 120
 DB 408 TTGAGCAAAAGCTTGTGATTAATCTCTCAGGTCTGCTTTCTGTACTTGCCTGGCCATTC 467
 QY 121 PheSerGlySerThrGlyAlaArgTyPheAlaAlaPheValProLeuValAsnGlyLeu 140
 DB 468 TTCAGCGGATGACCGAGCTCGATATCTTGTGCTTTGTTCGTTAGTAAGAGCTTA 527
 QY 141 ArgLeuValIleAsnGlyLeuSerIleSerProAnSerMetLeuIleLeuSerValThr 160
 DB 528 AGGCTGTATTATTAAGGACTATTCATTTCCCAATTTGAGTCTATCAATTCCTCA 587
 QY 161 ArgGlyArgGlyAlaGlyGlnLeuLeuLeuGlyProLeuPheTyValLeuAlaLeuLeu 180

DB 588 AGGAAGGGAGACAGAGGTGCTTAAGGTCCTTGTCTTACGTTACGTTACCTTCTT 647
 QY 181 PheSerAlaValPhePheTPARGLuSerProLLeuNHISGLInLeuCYAArgPheGlyAsnAsn 200
 DB 648 TTTCTGCGGATTTCTTCTGAGAGATCTTCTTATCGTATGATCTCTTATGCAATATG 707
 QY 201 CyGlyGlyAspGlyIleAlaAspIleMetGlyArgLySAspPheGlySerThrLySlePro 220
 DB 708 TGTGTGTGCGATGGAATAGCTGATATATATGAGACGTAAAGTTGGGTCAATAGATACCT 767
 QY 221 TyRAnProArgLySserTPAlaGlySerIleSerMetPheIlePheGlyPhePheIle 240
 DB 768 TACACACCAAGAAAGAGTTGGGACGAGAGCATCTCCATGTTCACTTCGCGCTTTCATC 827
 QY 241 SerIleAlaLeuLeuTyTyTyTySerSerLeuGlyTyRLeuHISMetAnTrpGlyThr 260
 DB 828 TCCATCGCATTTACTTACTTACTTCAAGCGCTTGGTACCTTCACTTAAGTGGAAACG 887
 QY 261 ThrLeuGlnArgValAlaMetValSerMetValAlaThrValAlaGlySerLeuProIle 280
 DB 888 ACCTTGACAGAGTATGCAATATGCTCAATGTAGCCACGGTGGTCGAGTGCATCCATC 947
 QY 281 ThrAspGlnLeuAspAspAnIleSerValProLeuAlaThrIleLeuAlaAlaTyLeu 300
 DB 948 ACCGATCAATTAGACGACGATTCGTTGCTCTCTGCTACTATTTAACTGCTTATTTA 1007
 QY 301 SerPheGlyTyR 304
 DB 1008 AGTTTCGGAATAT 1019

RESULT 4

ATT32M21

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

JOURNAL

TITLE

COMMENT

FEATURES

SOURCE

misc_feature

gene

/gene="T32M21_10"

/note="this clone overlaps with BAC T19M18"

/ecotype="Columbia"

/chromosome="5"

/db_xref="taxon:3702"

/mol_type="genomic DNA"

/organism="Arabidopsis thaliana"

location/Qualifiers

1..93045

2 (bases 1 to 93045)

Unpublished

Rudd,S., Lemcke,K. and Mayer,K.F.X.

villarreal,R., Gielen,J., Van Montagu,M., Bancroft,I., Mewes,H.W.,

Clerck,R., De Keyser,A., Neyt,P., Rouze,P., Van den Daele,H.,

Beyan,M., Terry,N., Ardiles,W., Buyschaert,C., Dasseville,R., De

1 (bases 1 to 93045)

rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;

Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Arabidopsis thaliana (thale cress)

AL162875.1 GI:7406444

ATT32M21

Arabidopsis thaliana DNA chromosome 5, BAC clone T32M21 (BSSA

project).

Submitted (31-MAR-2000) MIRP, at the Max-Planck-Institut fuer

Biochemie, Am Klopferspitz 18a, D-82152 Martinsried, FRG, E-mail:

leemcke@mps.biochem.mpg.de, mayer@mps.biochem.mpg.de Project

Coordinator: Mike Bevan, Molecular Genetics Department, Cambridge

Laboratory, John Innes Centre, Colney Lane, NR4 7UJ Norwich, UK,

E-mail: michael.bevan@brc.ac.uk

Information on performance of analysis and a more detailed

annotation of this entry and other sequences of chromosomes 3, 4

and 5 can be viewed at: <http://www.mips.biochem.mpg.de/proj/thal/>.

CDS
 join(676, .841,958, .1235,1354, .1545,1884, .2447,2525, .2642, 2727, .3112)
 /gene="TJ32M21_10"
 /note="strong similarity to NAC2, Arabidopsis thaliana, EMBL:AF201456"
 /codon_start=1
 /product="putative protein"
 /protein_id="CA85547.1"
 /db_xref="GI:7406445"
 /db_xref="GOA:Q84K00"
 /db_xref="UniProt/Swiss-Prot:Q84K00"
 /translation="MGKSVTSLAPGRPHPTDEELVRYYLKRRVCNKKPKPFAISVT
 DYKSEPMDLPKSKSLKSRDLFWYFPMILDKVYNSGSKTMRATEKGYWKTGQREK
 NGRVYGMKTYLHYHKGRAPRGERTMVMHEHYLSDEDLKAGVPGPAAVYLCI FQKS
 GTGPKNGEQGAPLEERWEDQNTVVPADAFSEGLALNDVYDIDDIIDEXENLY
 VYDAVPLPNYCHGESNNVSGNYSNGYIQGNVNVDSGGFEPQIPIDPHNYLDEP
 IIRGSIQPCSLFPEEQIGGVODENVNLSSNNVNVADTCYSIDIPIDHNYLDEP
 FMDPNNNLPLNDGILYLETNDLSCAQODFNFPEDYLSFPDEGLTPDSSLMGPEDFLP
 NORALDOKRAPKELKEKAGKAVKEKESGSSSKQDTPDPSAPKYPPLKXTS
 HMLGALTPSSFPASQFOTKDMRLHAAOSSGSHVATIGMMRISMTLADSGMGWSTD
 KNGNLNVLSFGVQDDAMTASGSKTGITATAMLVFCLWLLLSVSPKIVTWASA
 R"
 exon
 /gene="TJ32M21_10"
 /number=1
 676, .841
 intron
 /gene="TJ32M21_10"
 /number=1
 842, .957
 exon
 /gene="TJ32M21_10"
 /number=1
 958, .1235
 intron
 /gene="TJ32M21_10"
 /number=2
 1236, .1353
 intron
 /gene="TJ32M21_10"
 /number=2
 1354, .1545
 exon
 /gene="TJ32M21_10"
 /number=3
 1546, .1883
 intron
 /gene="TJ32M21_10"
 /number=3
 1884, .2447
 exon
 /gene="TJ32M21_10"
 /number=4
 2448, .2524
 intron
 /gene="TJ32M21_10"
 /number=4
 2525, .2642
 exon
 /gene="TJ32M21_10"
 /number=5
 2643, .2726
 intron
 /gene="TJ32M21_10"
 /number=5
 2727, .3112
 exon
 /gene="TJ32M21_10"
 /number=6
 3113, .6151
 gene
 /gene="TJ32M21_20"
 complement (3563, .3721,3811, .3870,3949, .4269, 4349, .4432,4513, .4601,4709, .4823,4900, .5026,5142, .5281, 5368, .5429,5507, .5573,5652, .5783,5870, .5947,6041, .6151)
 /gene="TJ32M21_20"
 complement (join(3563, .3721,3811, .3870,3949, .4269, 4349, .4432,4513, .4601,4709, .4823,4900, .5026,5142, .5281, 5368, .5429,5507, .5573,5652, .5783,5870, .5947,6041, .6151))
 /gene="TJ32M21_20"
 /note="similarity to rmgB protein, Dictyostellium discoideum, PIR:S68824"
 /codon_start=1
 /product="putative protein"
 /protein_id="CA85548.1"
 /db_xref="GI:7406446"
 /db_xref="InterPro:IPR006652"

/db_xref="InterPro:IPR011498"
 /db_xref="UniProt/TrEMBL:O91Z83"
 /translation="MEIRKEIDVGMHSHLADENWTPPLVSGSRASARYKHAAYVD
 EKLYIVGSGNRGLYSVDVQFDIRSLTWSLKTSSSDNLTQEDGSLREAFPAI
 SDHRMTKMGKRLILIGHSKSSDNNLVRFIDLETSCGTVDFGVNVPASRGHSITL
 VGSRLVFGGEGDKRRLLNDLHVLHLETMTDVTETKQTPVPRPHTAATHSDRYLL
 IFGGCSHSIFYSDLHILDIQTMESQPHVQGVVTPRAGAGITIDENWYIVGCGNS
 TGCLETVLVMSKLTWSTHTVBARHPLASBGLSVCSASVFGENILVAFGVNGKN
 DIFWMLKRGSSHPKIFKSPAAAAAATYAAATAKSKSPPPRAPATTAGNCS
 LPFRDIRNRIDTKEERLLESSIAETQVENALRKRIDVNSHTLSGELSVBGO
 LISERRCFPLEAQLAELOKLESQSIEAEVAMLRQRASDEEDGTVOQGSAGV
 MGLFGR"
 exon
 complement (3563, .3721)
 /gene="TJ32M21_20"
 /number=1
 intron
 complement (3722, .3810)
 /gene="TJ32M21_20"
 /number=1
 exon
 complement (3811, .3870)
 /gene="TJ32M21_20"
 /number=2
 intron
 complement (3871, .3948)
 /gene="TJ32M21_20"
 /number=2
 exon
 complement (3949, .4269)
 /gene="TJ32M21_20"
 /number=3
 intron
 complement (4270, .4348)
 /gene="TJ32M21_20"
 /number=3
 exon
 complement (4349, .4432)
 /gene="TJ32M21_20"
 /number=4
 intron
 complement (4433, .4512)
 /gene="TJ32M21_20"
 /number=4
 exon
 complement (4513, .4601)
 /gene="TJ32M21_20"
 /number=5
 intron
 complement (4602, .4708)
 /gene="TJ32M21_20"
 /number=5
 exon
 complement (4709, .4823)
 /gene="TJ32M21_20"
 /number=6
 intron
 complement (4824, .4899)
 /gene="TJ32M21_20"
 /number=6
 exon
 complement (4900, .5026)
 /gene="TJ32M21_20"
 /number=7
 intron
 complement (5027, .5141)
 /gene="TJ32M21_20"
 /number=7
 exon
 complement (5142, .5281)
 /gene="TJ32M21_20"
 /number=8
 intron
 complement (5282, .5367)
 /gene="TJ32M21_20"
 /number=8
 exon
 complement (5368, .5429)
 /gene="TJ32M21_20"
 /number=9
 intron
 complement (5430, .5506)
 /gene="TJ32M21_20"
 /number=9
 exon
 complement (5507, .5573)
 /gene="TJ32M21_20"
 /number=10
 intron
 complement (5574, .5651)
 /gene="TJ32M21_20"
 /number=10
 exon
 complement (5652, .5783)

/gene="T32M21_20"
 /number=11
 complement (5784, .5869)
 /gene="T32M21_20"
 /number=11
 complement (5870, .5947)
 /gene="T32M21_20"
 /number=12
 complement (5948, .6040)
 /gene="T32M21_20"
 /number=12
 complement (6041, .6151)
 /gene="T32M21_20"
 /number=13

Alignment Scores:

Pred. No.: 6,08e-105 Length: 93045
 Score: 1299.50 Matches: 302
 Percent Similarity: 52.6% Conservaive: 0
 Best Local Similarity: 52.6% Mismatches: 2
 Query Match: 84.6% Indels: 272
 DB: 15 Gaps: 5

US-10-634-548-2 (1-304) x ATT32M21 (1-93045)

QY 1 MetAlaAlaThrLeuProLeuSerProIleAsnHsgInLeuCy8ArgPheGlyAsnAsn 20
 Db 36563 ATGGCAGCAACCTTAACCTCTATCCGATCATCATCAGTTGTGTGGGAAACAC 36622
 QY 21 SerLeuThrThrHsaArgPheCy8SerProGlyPheLeuIleSerSerProCy8PheIle 40
 Db 36623 TCTTTCAGACTCCACCGGCTTCGTCTCCGCTTCTGATTTCTTCTCTGTTTCAAT 36682
 QY 41 GlyLeuThrGlyMetGlySerAlaThrGlnLeuArgAlaArgSerLeuIleSerSer 60
 Db 36683 GGTTCGACCGGAATGGGCTCTGCTACTACGTTACGTCGTCTCTCTATCTCTTCA 36742
 QY 61 AlaValAlaThrAsnSerLeuLeuHsaPylGlyAlaThrValAlaValLeuGlyGly 80
 Db 36743 GCAGTTGGACGAATTCGCTGTGTCATGACGTCGAGCCACCGTGGCAGTGTGGTGA 36802
 QY 81 AlaThrAlaLeuValLeuSerPheGlySerLeuThrIlyAsnArgAsnValIleGln----- 98
 Db 36803 GCATACGGCGCTTGTCTTAAGCTTCGAGAGTCTCAACCAAGCAACGTCATTCACACAGGTC 36862
 QY 98 ----- 98
 Db 36863 TCTTAATATATGTTTAACTTATCCACACAAATTTCTCCGTTTAACTTCCAGTTTATTTCG 36922
 QY 98 ----- 98
 Db 36923 AACACACTATATGTTGAAAAGATTCTCAAGTTGTGTTGCAGTACTACTATAGAA 36982
 QY 98 ----- 98
 Db 36983 ACAATGATATACCTAGAAATTTTGTGTGTAATAGTTTTCATTCTGAATTTTATTA 37042
 QY 98 ----- 98
 Db 37043 GAATTGGTAAACCTTAACTAGACATATACCACTTATCATGACCAATCGGTAAGCGGA 37102
 QY 98 ----- 98
 Db 37103 CAAGAACAAGTGTCCAAAAATATTACCGCTTATATATGTTACCACTTCTCTTAACCTC 37162
 QY 98 ----- 98
 Db 37163 CCTTTTAATATCCGTAAATCCGTACCGCTAAACATATACCGTTCCTTGTGTTAACA 37222
 QY 99 -----GlnSerLeuSerAr 103
 Db 37223 AAGTAAAGAAAGAAAGAAACATTAATTGATTTGTTTATGTGTGACGAGATTGAGCAG 37282

QY 103 GlyLeuValIleIleLeuSerGlyLeuLeuPheValLeuAlaTrpProfile----- 120
 Db 37283 AAGCTGTGCATATACTCTCAGGTCTGCTTTCGTACTTGGCGGCAATCTTCAGGTA 37342
 QY 120 ----- 120
 Db 37343 TTGCTTCTCTATATGTTTGTAAATCTCTGTACTTCTTAAACATATATGACTCTGA 37402
 QY 121 -----PheSerGlySerThrGlyAlaArgIlyrPheAlaIaph 133
 Db 37403 TTTCTTTTACTCATCTTAAAGTTAGCGGATGACCGAGGCTCGATCTTGTGCTTT 37462
 QY 133 eValProLeuValIleGlyLeuArgLeuValIleAsnGlyLeuSerIleSerProAsnSe 153
 Db 37463 TGTTCGCTAGTGAATGGCTTAAAGGCTTGTATTAACGACATATCCATTTCCCAAAATTC 37522
 QY 153 rMetLeuIleLysSerValThrArgGlyLysIlyrArgAlaGlu----- 166
 Db 37523 GATGCTAATCAAAATCCGTCAACAAGAGAGAGAGAGAGTAAAGTGTCTAATTTTTC 37582
 QY 166 ----- 166
 Db 37583 TCCAACTTGATATGATTTTCAACAATCTGATTACATTTCTTCTTTCACACATCA 37642
 QY 167 -----GlnLeuLeuLysGlyProLeuPheIlyrValLeuAlaLeuLeuPheSerAlaValPh 185
 Db 37643 CAGAGAGTGTCTTAAAGTCTTGTGTTCTACGTTCTACCTTCTTCTCTGCGGTTTT 37702
 QY 185 ePheThrArgGlySerProIleGlyMetIleSerLeuAlaMetMetCy8GlyGlyAsnGly 205
 Db 37703 CTTCGAGAGAGTCTCTATTCGGTATGATCTGTTAGCAATGATATGTGTGGCGATGG 37762
 QY 205 yIle----- 206
 Db 37763 TAA-ATTTTCGCAAGTACTACTGATTAATCTATTAACAATTACAAATGGCATAA 37821
 QY 207 -----AlaAspIleMetGlyAr 212
 Db 37822 ATGACTAATCAATAGTCGATCATATATGTCATATGTAAGATATGATATTAATGGACG 37881
 QY 212 GlyPheGlySerThrIlyrIleProIlyrAsnProArgLysSerTrpAlaGlySerIleSe 232
 Db 37882 TAAAGTTGGGTCATTAAGATACCTTACACCCCAAGAAAGATTGGGACAGAACATCTC 37941
 QY 232 rMetPheIlePheGlyPhePheIleSerIleAla----- 243
 Db 37942 CATGTTATCTGGCTTCTTCAATCTCCATCGC-GTAAAAATATTACCAATCCCATATT 38000
 QY 243 ----- 243
 Db 38001 AATCATCAAAATGTCCTCTCTTGACGACGAACAAGCTTAAAGAACTGAGATGATTCG 38060
 QY 244 -----LeuLeuThrIlyrIlyrSerSerIe 251
 Db 38061 TACTAAACCTTAACGTTTCTTTTGTAAATTTTGACAGATTAATCTTACTATTAATCAAGCT 38120
 QY 251 uGlyIlyrLeuHsiMetAsnTrpGlyIlyrThrIleuGlnArgValAlaMetValSerMetVa 271
 Db 38121 TGAGTACCTTCACATGAATGGGAAACGACCTTCAGAGATGACCAATGCTCAATGCT 38180
 QY 271 lAlaThrValAlaGlySerLeuProIleThrAsnGlnLeuAspAsnIleSerValPr 291
 Db 38181 CGCCACCGTATGTCAGTGCCTACCAACCAACCGAATTAATGACCAATATTTGGTTCC 38240
 QY 291 oLeuAlaThrIleLeuAlaIlyrLeuSerPheGlyIlyr 304
 Db 38241 TCTGGCTACTATTTTAACTGCTTATTTAAAGTTTGGATAT 38280
 QY 38241 -----

RESULT 5
 AK061265 1002 bp mRNA linear PLN 24-JUL-2003
 LOCUS AK061265
 DEFINITION Oryza sativa (japonica cultivar-group) cDNA clone:006-212-C06, full
 insert sequence.


```
Db 26849 TTGGATTGTTCTGTTGCAATGATGAGTGGTGGATGTA- GATTTGATCTCTG 26907
Qy 206 ----- 206
Db 26908 GGAATTAGGAAATCTGACATTTTTTTTGTAAATACCACATTCACTTGATGAATGA 26967
Qy 207 ----- AlaaplleM 210
Db 26968 TGCCTTGCATAGATATACCTCTTGATTTTATTTTTCAGGCTTGTGACATG 27027
Qy 210 etgIATgLyPhegISeThrLySleProTyAsnProArgLySeTrpAlaGly 230
Db 27028 TTGGAGAGGATAGGCTCCGCGAAGCTGCATTCANAGACAGAGCTGATAGGA 27087
Qy 230 erIleSeMetPheIlePhegIyPhePheIleSerIleAla----- 243
Db 27088 GCATCTCAATGTCATTTCTGCTTCTTATCTGC- ACTGTAGTTTTTGTTCAGAA 27146
Qy 243 ----- 243
Db 27147 TGCATTGCTGTTCTTTTGCATTGTACTCTTGTATTTGTAATGATCTTTTGC 27206
Qy 244 -----LeuLeuTyTyrTySerserLeuGlyTyLeuHiMetAsnTrpLuhT 261
Db 27207 ATCAAGAGATGCTTCTTACTCTTCTGCTTGTACTTACAGTCTGCTGGATCTGG 27266
Qy 261 hrIeugIAnrGvAlAlMeValseMetValAlathrValAlgluSerLeuProIet 281
Db 27267 CACTTGATGATGCTCTTGTGCTTGTGATGAGCTACTGATGAGTATCTCTGCA 27326
Qy 281 hrIeugIAnrGvAlAlMeValseMetValAlathrValAlgluSerLeuProIet 301
Db 27327 ATGATGTGAGATGACATATTCCTCTTGGCCACATGTGACACCTATCTGT 27386
Qy 301 erPhegIyTyr 304
Db 27387 TATTGGCTAC 27397
```

RESULT 7
AP008210_342
WPCOMMENT

Sequence split into 355 fragments. LOCUS AP008210 Accession AP008210

Fragment Name	Begin	End
AP008210_000	1	110000
AP008210_001	100001	210000
AP008210_002	200001	310000
AP008210_003	300001	410000
AP008210_004	400001	510000
AP008210_005	500001	610000
AP008210_006	600001	710000
AP008210_007	700001	810000
AP008210_008	800001	910000
AP008210_009	900001	1010000
AP008210_010	1000001	1110000
AP008210_011	1100001	1210000
AP008210_012	1200001	1310000
AP008210_013	1300001	1410000
AP008210_014	1400001	1510000
AP008210_015	1500001	1610000
AP008210_016	1600001	1710000
AP008210_017	1700001	1810000
AP008210_018	1800001	1910000
AP008210_019	1900001	2010000
AP008210_020	2000001	2110000
AP008210_021	2100001	2210000
AP008210_022	2200001	2310000
AP008210_023	2300001	2410000
AP008210_024	2400001	2510000
AP008210_025	2500001	2610000
AP008210_026	2600001	2710000
AP008210_027	2700001	2810000
AP008210_028	2800001	2910000

AP008210_029	2900001	3010000
AP008210_030	3000001	3110000
AP008210_031	3100001	3210000
AP008210_032	3200001	3310000
AP008210_033	3300001	3410000
AP008210_034	3400001	3510000
AP008210_035	3500001	3610000
AP008210_036	3600001	3710000
AP008210_037	3700001	3810000
AP008210_038	3800001	3910000
AP008210_039	3900001	4010000
AP008210_040	4000001	4110000
AP008210_041	4100001	4210000
AP008210_042	4200001	4310000
AP008210_043	4300001	4410000
AP008210_044	4400001	4510000
AP008210_045	4500001	4610000
AP008210_046	4600001	4710000
AP008210_047	4700001	4810000
AP008210_048	4800001	4910000
AP008210_049	4900001	5010000
AP008210_050	5000001	5110000
AP008210_051	5100001	5210000
AP008210_052	5200001	5310000
AP008210_053	5300001	5410000
AP008210_054	5400001	5510000
AP008210_055	5500001	5610000
AP008210_056	5600001	5710000
AP008210_057	5700001	5810000
AP008210_058	5800001	5910000
AP008210_059	5900001	6010000
AP008210_060	6000001	6110000
AP008210_061	6100001	6210000
AP008210_062	6200001	6310000
AP008210_063	6300001	6410000
AP008210_064	6400001	6510000
AP008210_065	6500001	6610000
AP008210_066	6600001	6710000
AP008210_067	6700001	6810000
AP008210_068	6800001	6910000
AP008210_069	6900001	7010000
AP008210_070	7000001	7110000
AP008210_071	7100001	7210000
AP008210_072	7200001	7310000
AP008210_073	7300001	7410000
AP008210_074	7400001	7510000
AP008210_075	7500001	7610000
AP008210_076	7600001	7710000
AP008210_077	7700001	7810000
AP008210_078	7800001	7910000
AP008210_079	7900001	8010000
AP008210_080	8000001	8110000
AP008210_081	8100001	8210000
AP008210_082	8200001	8310000
AP008210_083	8300001	8410000
AP008210_084	8400001	8510000
AP008210_085	8500001	8610000
AP008210_086	8600001	8710000
AP008210_087	8700001	8810000
AP008210_088	8800001	8910000
AP008210_089	8900001	9010000
AP008210_090	9000001	9110000
AP008210_091	9100001	9210000
AP008210_092	9200001	9310000
AP008210_093	9300001	9410000
AP008210_094	9400001	9510000
AP008210_095	9500001	9610000
AP008210_096	9600001	9710000
AP008210_097	9700001	9810000
AP008210_098	9800001	9910000
AP008210_099	9900001	10010000
AP008210_100	10000001	10110000
AP008210_101	10100001	10210000

[illegible]

Db 44774 TATTGGCTAC 44784

RESULT 8
OSUN00032

LOCUS
DEFINITION
Oryza sativa genomic DNA, chromosome 4, BAC clone: OSUNBA0043A12,
complete sequence.

ACCESSION
AL606619
VERSION
GI:32488370

SOURCE
ORGANISM
Oryza sativa {japonica cultivar-group}
HTG.
Oryza sativa {japonica cultivar-group}

REFERENCE
AUTHORS
Feng Q., Zhang Y., Hao P., Wang S., Fu G., Huang Y., Li Y., Zhu J.,
Liu Y., Hu X., Jia P., Zhang Y., Zhao Q., Ying K., Yu S., Tang Y.,
Weng Q., Zhang L., Lu Y., Mu J., Lu Y., Zhang L.S., Yu Z., Fan D.,
Lu X., Lu T., Li C., Wu Y., Sun T., Lei H., Li T., Hu H., Guan J.,
Wu M., Zhang R., Zhou B., Chen Z., Chen L., Jin Z., Wang R.,
Yin H., Cai Z., Ren S., Lv G., Gu W., Zhu G., Tu Y., Jia J.,
Zhang Y., Chen J., Kang H., Chen X., Shao C., Sun Y., Hu O.,
Zhang X., Zhang W., Wang L., Ding C., Sheng H., Gu J., Chen S.,
Ni L., Zhu F., Xue W., Lan L., Lai Y., Cheng Z., Gu M., Jiang J.,
Li J., Hong G., Chen Y. and Han B.
Sequence and analysis of rice chromosome 4
Nature 420 (6913), 316-320 (2002)

JOURNAL
PUBMED
12447439

TITLE
REFERENCE
AUTHORS
Fu G., Wang S.Y., Ren S.X., Lv G., Lin W., Gu W.Q., Zhu G.F.,
Tu Y.F., Jia J., Yin H.F., Zhang Y., Cai Z., Chen J., Kang H.,
Chen X.Y., Shao Y., Sun Y., Hu Q.P., Zhang X.L., Zhang W.,
Wang L.J., Ding C.W., Sheng H.H., Gu J.L., Chen S.T., Ni L.,
Zhu F.H., Han B., Feng Q., Huang Y.C., Li Y., Zhu J.J., Zhao Q.,
Hu X., Liu Y.L., Mu J., Yu Z., Chen L., Fan D.L., Weng Q.J.,
Zhang L., Lu Y., Yu S.L., Liu X.H., Lu T.T., Zhang Y.J., Lu Y.,
Li C., Li T., Zhang Y., Hu H., Jia P.X., Qian Y.M., Ying K.,
Zhou B., Chen Z.H., Hao P., Zhang L., Wu M., Zhang R.Q., Guan J.P.
and Hong G.F.

TITLE
JOURNAL
COMMENT
REMARK
SOURCE

Direct Submission
Submitted (08-SEP-2001) Han Bin, National Center for Gene Research,
Chinese Academy of sciences, 500# Cao Bao Road, Shanghai 200233,
CHINA. E-mail enquiries: bhan@ncgr.ac.cn. Clone requests:
bhan@ncgr.ac.cn
Oryza sativa japonica (niponbare) genomic DNA, chromosome 4, BAC
clone: OSUNBA0043A12.
On Jul 9, 2003 this sequence version replaced gi:21912468.
Web site: http://www.ncgr.ac.cn
Summary Statistics
Assembly program: phrap

This is a complete sequence.
Genes were identified by a combination of several methods: Gene
prediction programs including Genesh (<http://www.softberry.com/>),
genescan (<http://CCR-081.mit.edu/GENSECAN.html>), GenemarkRHM
(<http://genemark.biology.gatech.edu/genemark/>), tRNAscan-SE (Sean
Eddy, <http://genome.wustl.edu/eddy/tRNAscan-SE/>), searches of the
complete sequence against NCBI none redundant protein database (nr)
(<ftp://ncbi.nlm.nih.gov/blast/db>) and the EST database at NCGC.
Location:Nullifiers
1. 190432
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="genomic DNA"
/cultivar="japonica"
/cultivar="Niponbare"
/sub_species="japonica"
/db_xref="taxon:39947"
/chromosome="4"
/clone="OSUNBA0043A12"
/clone_11b="CUIG-OSUNBA"
.7612

gene


```

Qy 19 AsnAsnSerLeu-----ThrThiAlaPheCysSerProGlyPheLeuIle 34
Db 154547 TCTCTCTCTCTGCTGCTGCGCGCTCCAGTCCGCGCTCTCTGCGCGCGCGCG 154606
Qy 35 SerSer-ProGlyPheIleGlyLeu--ThrGlyMetGlySerAlaThrGlnLeuAla 53
Db 154607 GCGAGCTCCATGCGCGCGAGGCTTGTGCTGCGCGCGCGCGCGCGCGCGCGCG 154657
Qy 54 ArgArgSerLeuIleSerSerAlaAlaAlaThrAsnSerLeuLeuH1sAspValGlyAla 73
Db 154658 CCGGCGCTGCGCGCGCTGCGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCG 154717
Qy 74 ThrValAlaValLeuGlyGlyAlaAlaThrAlaLeuValLeuSerPheGlySerLeuThrLys 93
Db 154718 ACCTGCTCATCACCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 154777
Qy 94 ArgAsnValIleGln----- 98
Db 154778 CCGCGGCTCATGCAACAGGCTGCGGATCTGTTTCTTTTCTTTTCTTTTCTTTTCTTTT 154837
Qy 98 ----- 98
Db 154838 TTTTGAAGTGAAGTGAAGTCTTGTATCTCGGCATCTTCAAAAATATCATCTTTAG 154897
Qy 98 ----- 98
Db 154898 CTCTGTTGACTGTTATCCGAGAGTAATATGATATATCATATTCAGTGAATGATCT 154957
Qy 98 ----- 98
Db 154958 GTATCTGAATTTATTTTCAATTGCTGGAAGATCTAGTGTTCAGATGAGCTGTAAT 155017
Qy 98 ----- 98
Db 155018 GTTCAAAAATACAGTAGTAGATAGATATAAAGAGAGTCTTCTTATATGATCATCA 155077
Qy 98 ----- 98
Db 155078 AAGAAAATAGATTATTCAGCATMAAATTATACCTTTCTTATATATCTTGTTAAT 155137
Qy 99 ----- 99
Db 155138 TCTGATCTTTTCAATATATATCTCAGACCTGAGCAAAAATATGATGCTGCTCTCC 155197
Qy 111 GlyLeuPheValLeuAlaAlaPheProIlePhe----- 121
Db 155198 GAGATCTGTTCAATGCTCTGCGCATATTCAGTCTCCCTTGCACATGCTTCTTCCA 155257
Qy 121 ----- 121
Db 155258 GCATCGAATGTTTCTTCAACAACCAATCAGATAGATCTTAAAGAAATCAGCCCTT 155317
Qy 122 SerGlySerThrGlnAlaArgTyrPheAlaAlaPheValProLeuValaGlyLe 140
Db 155318 TCGCAGTAATTCAGACAGAGACGCTTCTTCCGCGCATATGTCCTGCTGCAATGCAAT 155377
Qy 140 ValGlyLeuValIleAlaGlyLeuSerIleSerProAsnSerMetLeuIleLysSerValTh 160
Db 155378 AAGGCTTCTGACCTACGCGGCGCTCCCTTCCACATGAGAAAGCTTATGAAATAGGAGAC 155437
Qy 160 PArgGlyGlyArgAlaGly----- 166
Db 155438 CCGTGAAGGAAACACAGAGTTCGTATGATCTTATTTCTGATACCTCTTCAATGAAA 155497
Qy 167 ----- 172
Db 155498 AATCAAAATAGTATCTTCAATGATCTTGTGCTTACAGGAAATGCTGAGAGCTC 155557
Qy 172 roleupheryValLeuAlaLeuLeuPheSerAlaValPhePheTyrArgGlySerProI 192
Db 155558 CTCTGTAATATGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 155617

```

```

Qy 192 leglyMetIleSerLeuAlaMetMetCysGlyGlyAspGlyIle----- 206
Db 155618 TTGGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 155676
Qy 206 ----- 206
Db 155677 GGAATTTAGAAATATGACAGTTTCTTTTGTATATACCCAGTTCACTTATGATGAAATGAA 155736
Qy 207 ----- 210
Db 155737 TGCTTCTGTAAGATATATCCACTTCTTGTGTTCTTATTTTTCAGGCTTGTGCACTTG 155796
Qy 210 etGlyArgLysPheGlySerThrLysIleProTyrAsnProArgLysSerTyrAlaLys 230
Db 155797 TTGGAGAGAGTATGCTTCCGCGAGCTGCTCATATGAGAAACAGAGCTGATAGGAA 155856
Qy 230 erIleSerMetPheIlePheGlyPhePheIleSerIleAla----- 243
Db 155857 GCATCTCAATGTTCAATTTCTGCTGCTCTCTATCTGC-ACGTAGAGTTTGTGCTTACAGAA 155915
Qy 243 ----- 243
Db 155916 TGCAATGCTGTTCTTTTGCATTTGTTACTCTTGTATTTGTTAATCTGATGATCTTTGC 155975
Qy 244 -----LeuLeuTyrTyrTyrSerSerLeuGlyTyrLeuH1sMetAsnTyrGlnThr 261
Db 155976 ATCAAGAGATGCTGTTTACTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 156035
Qy 261 hrLeuGlnArgValAlaMetValSerMetValAlaThrValAlaGlySerLeuProIleT 281
Db 156036 CACTTGTAACAGTGGCTCTGTTGATACAGCTATGTAAGTGAATGATATCTGCTCA 156095
Qy 281 hrAspGlnLeuAspAspAsnIleSerValProLeuAlaThrIleLeuAlaAlaTyrLeu 301
Db 156096 ATGATGTTGGATGACAAATATCTCCCTTCTTGGCCACAGATTTGCGAGCTTATCTGT 156155
Qy 301 erPheGlyTyr 304
Db 156156 TATTTGGCTAC 156166

```

RESULT 9
AK110748
LOCUS
DEFINITION
Oryza sativa (japonica cultivar-group) cDNA clone:002-170-G04, full insert sequence.
AK110748
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE
AUTHORS

The Rice Full-length cDNA Consortium, National Institute of Agricultural Sciences Rice Full-length cDNA Project Team, Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T., Foundation of Advancement of International Science Genome Sequencing & Analysis Group, Otono, Y., Murakami, K., Iida, Y., Sugano, S., Fujimura, T., Suzuki, Y., Tanoda, Y., Kurosaki, T., Kodama, T., Masuda, H., Kobayashi, M., Xie, Q., Lu, M., Narikawa, R., Sugiyama, A., Mizuno, K., Yokomizo, S., Nikiura, J., Ikeda, R., Ishibiki, J., Kawamata, M., Yoshimura, A., Miura, J., Kusumegi, T., Oka, M., Ryu, R., Ueda, M., Matubara, A., Kikawa, T., Kawai, J., Carninci, P., Adachi, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashidume, W., Hayatsu, N., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kondo, S., Kono, H., Miyazaki, A., Otsu, N., Ota, Y., Saito, R., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Yoshino, M., and Hayashizaki, Y.
Collection, mapping, and annotation of over 28,000 cDNA clones from japonica rice

JOURNAL Science 301 (5631), 376-379 (2003)
 PUBMED 12869764
 REFERENCE 2 (bases 1 to 1218)
 AUTHORS Adachi, J., Aizawa, K., Akimura, T., Arikawa, T., Carninci, P., Doi, K., Fujimura, T., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayashizaki, Y., Hayatsu, N., Hirooka, T., Hirooka, T., Hori, P., Hotta, I., Iida, J., Iida, Y., Ikeda, R., Imamura, K., Imotani, K., Ishibiki, J., Ishii, Y., Ishikawa, M., Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawagashira, N., Kawai, J., Kawamata, M., Kikuchi, S., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M., Kodama, T., Kojima, K., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Kurosaki, T., Kusumegi, T., Li, C., Lu, M., Maeda, H., Matsubara, K., Matsuyama, T., Miura, J., Miyazaki, A., Mizuno, K., Murakami, K., Murata, M., Nagata, T., Nakamura, M., Namiki, T., Narikawa, R., Nishikawa, J., Nishii, K., Nomura, K., Numasaki, R., Ohneda, E., Ohno, M., Ohsuki, K., Oka, M., Ooka, H., Oshino, N., Ota, Y., Otsu, Y., Otsu, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Sato, K., Shibata, K., Shingawa, A., Shiraki, T., Shishiki, T., Sugano, S., Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Tsunoda, Y., Ueda, M., Waki, K., Xie, Q., Yanagi, W., Yamada, H., Yamamoto, M., Yasunishi, A., Yazaki, J., Yokomizo, S. and Yoshimura, A.

Direct Submission
 Submitted (27-AUG-2002) Shoshi Kikuchi, National Institute of Agricultural Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression, 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602, Japan (E-mail:skikuchi@nias.affrc.go.jp).
 Tel:81-29-838-7007, Fax:81-29-838-7007
 This clone is one of the 28K full-length cDNA clones from japonica rice.

URL: http://cdna01.dna.affrc.go.jp/cDNA/NIAS_Rice_Full-length_cDNA_Project_Team_Kikuchi_S_Sato_K_Nagata_T_Kawagashira_N_Doi_K_Kishimoto_N_Yazaki_J_Ishikawa_M_Yamada_H_Ooka_H_Hotta_I_Kojima_K_Namiki_T_Ohneda_E_Yanagi_W_Suzuki_K_Li_C_Ohsuki_K_Shishiki_T_Yamamoto_M

PLS Genome Sequencing & Analysis Group: Otsu, Y., Iida, Y., Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M., Kobayashi, M., Kodama, T., Kurosaki, T., Kusumegi, T., Lu, M., Maeda, H., Miura, J., Mizuno, K., Narikawa, R., Nishikawa, J., Oka, M., Otsu, R., Sugano, S., Sugiyama, A., Suzuki, Y., Tsunoda, Y., Ueda, M., Xie, Q., Yokomizo, S., Yoshimura, A., Matsubara, K. and Murakami, K.

Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K., Akimura, T., Arikawa, T., Carninci, P., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hirooka, T., Hirooka, T., Hori, F., Iida, J., Imamura, K., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawai, J., Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishii, K., Nomura, K., Numasaki, R., Ohno, M., Oota, N., Ota, Y., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Shibata, K., Shingawa, A., Shiraki, T., Sugawara, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Waki, K., Yasunishi, A. and Hayashizaki, Y.

Location/Qualifiers
 1..1218
 /organism="Oryza sativa (japonica cultivar-group)"
 /mol_type="mRNA"
 /cultiyar="Nipponbare"
 /db_xref="taxon:39947"
 /clone="002-170-G04"

ORIGIN

Alignment Scores:

Pred. No.: 1,11e-38 Length: 1218
 Score: 536.50 Matches: 123
 Percent Similarity: 59.0% Conservative: 48
 Best Local Similarity: 42.4% Mismatches: 100
 Query Match: 34.9% Indels: 19

DB: US-10-634-548-2 (1-304) x AK110748 (1-1218) 15 Gaps: 8

Qy 18 G|YAsnAsnSerLeuThrTrhHsArgPheCysSerProGlyPheLeuIleSerSerPro 37
 Db 101 GGAATTAATCA-----AGTCACGACACCC-----TTAATATCATACAGTA 142
 Qy 38 CyPheIleGlyLeuThrGlyMetGlySerAlaThrGlnLeuHrAlaArgSerLeu 57
 Db 143 TGTAAATCTGAGATTGTGTC-----CGCTCCGCTAGCAAGCC 181
 Qy 58 ILeSerSerAlaValAlaThrAsnSerLeuHsAspValGlyAlaThrValAlaVal 77
 Db 182 TCAATCATGATGATTCGAGAGCCGCTGTCGAGACCCGAGACCCGCGCC-----GCCGTG 235
 Qy 78 LeuGlyGlyAlaThrAlaLeu---ValLeuSerPhe---GluSerLeuThrTrhArgAsn 95
 Db 236 CTCACCGAGCCACCGCCCTGCGCTTCGTTCTGGAGAGAGTCCGACACCGCCGCC 295
 Qy 96 ValIleGlnGlnSerLeuSerArgGlyLeuValHsIleLeuSerGlyLeuLeuPheVal 115
 Db 296 CTCCTTGACCAAGAACTCTGCAGAAACTCTGCATTTACTCTGCTTACTTACTTC 355
 Qy 116 LeuAlaTrpProIlePheSerGly---SerThrGlnAlaArgTrpPheAlaAlaPheVal 134
 Db 356 CTATATGCGCTTGTTCAGTGCAGATGATGATCTCATTTCTGCTTCTCCCTGTC 415
 Qy 135 ProLeuValAsnGlyLeuArgLeuValIleAsnGlyLeuSerIleSerProAsnSerMet 154
 Db 416 ATTCATTCACATCAATCAATGATTCCTGATTTGAGTTCGATTTGATTCATCAAGATGATG 475
 Qy 155 LeuIleLeuSerValThrArgGlnGlyArgAlaGlnGlnLeuLeuGlyGlyProLeuPhe 174
 Db 476 GTCTATTACTCTCATGACGAGAAAGAGAGACGAGAGAGCTTCTTAAGTCCATCTTAC 535
 Qy 175 TyrValLeuAlaLeuLeuPheSerAlaValPhePheTrpArgGlnSerProIleGlyMet 194
 Db 536 TAGCGTTGCTTAAACATCGCTACTGTAATTTCTGGAGAGCATCTCCATTCATTCATC 595
 Qy 195 ILeSerLeuAlaMetMetCysGlyGlyArgGlyIleLeuIleAsnGlyArgGlyPhe 214
 Db 596 GCAAGTATTTGCAACTGTGTGCGGAGATGTGTACTCAATCAATGCTGAAAGCGCATTC 655
 Qy 215 GlySerThrTrpIleProIleProIleProIleProIleProIleProIleProIlePro 234
 Db 656 GGCATGTAATGCTTCTTCAACCTGATGATGATGATGATGATGATGATGATGATGATG 715
 Qy 235 ILePheGlyPhePheIleSerIleAlaLeuLeuTrpTrpTrpSerSerLeuGlyTyrLeu 254
 Db 716 TTAGCTGTTCCTTGTGATCAATCTGTACATGCTGATGCTGATGCTGATGCTGATGCT 775
 Qy 255 HsMetSerTrpGlnTrpTrhLeuGlnArgValAlaMetValSerMetValAlaThrVal 274
 Db 776 GAGGAGAGCTGAGCATGATGATC---GCCTTGCTGTGACATGCTGCTGCTGCAATY 832
 Qy 275 ValGlnSerLeuProIleTrhAspGlnLeuAspAspAsnIleSerValProLeuAlaThr 294
 Db 833 GTGAGTCATCTCTTATCAGACACGCTGAGACACATCTGATGTGCTGCTGCTCT 892
 Qy 295 ILeLeuAlaAlaTyrLeuSerPheGlyTyr 304
 Db 893 GTCTCTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 922

RESULT 10
 AK176090 1211 bp mRNA linear PLN 09-SEP-2004
 LOCUS Arabidopsis thaliana mRNA, complete cds, clone: RAF122-86-H21.
 DEFINITION AK176090
 ACCESSION AK176090
 VERSION AK176090.1 GI:51970321
 KEYWORDS FLI CDNA.
 SOURCE Arabidopsis thaliana (thale cress)
 ORGANISM Arabidopsis thaliana

Aktiyama, K., Satou, M., Toyoda, T., Konagaya, A., Carninci, P.,
 Kawai, J., Hayashizaki, Y., and Shinozaki, K.
 Large-scale analysis of RIKEN Arabidopsis full-length (RAFL) cDNAs
 Unpublished
 2 (bases 1 to 1360)
 Totsuki, Y., Seki, M., Ishida, J., Nakajima, M., Enju, A., Kamiya, A.,
 Narusaka, M., Shin, I., Nakagawa, M., Sakamoto, N., Oishi, K.,
 Kohara, Y., Kobayashi, M., Toyoda, A., Sakaki, Y., Sakurai, T., Iida, K.,
 Aktiyama, K., Satou, M., Toyoda, T., Konagaya, A., Carninci, P.,
 Kawai, J., Hayashizaki, Y., and Shinozaki, K.
 Direct Submission
 Submitted (06-SEP-2004) Motoaki Seki, RIKEN Genomic Sciences
 Center, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa
 230-0045, Japan (E-mail: msek1@gsc.riken.jp,
 URL: http://range.gsc.riken.jp/, Tel: 81-45-503-9625,
 Fax: 81-45-503-9586)
 An Arabidopsis full-length cDNA library was constructed essentially
 as reported previously (Seki et al. (1998) Plant J. 15:707-720;
 Seki et al. (2002) Science 296:141-145). cDNA cleaved with BamHI
 and XhoI was ligated to modified lambda PhiC-1 vector (Carninci et
 al. (2001) Genomics 77:79-90) digested with BamHI and SalI. This
 clone is in a modified pBluescript vector.
 Please visit our web site (http://range.gsc.riken.jp/) for further
 details.

FEATURES
 source
 1..1360
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /db_xref="taxon:3702"
 /chromosome="5"
 /clone="RAFL23-09-006"
 /ecotype="Columbia"
 /note="Common name: thale cress"
 1..1360
 /gene="At5g58560"
 120..1043
 /gene="At5g58560"
 /codon_start=1
 /product="unknown protein"
 /protein_id="BA043980.1"
 /db_xref="GI:51970576"
 /translation="MATTSTTKKSLVLCSEFSSPLVDSPSLAFSPPIPELTARIA
 TSPSSSRFPATKIRKSLAAMPENSVSDVCAFVTSIVAFSGFGEIGKGI
 FDOELIRLVINIGLVMLCPFSSGIGALPASLPGINIVRMILGIGVYHDEG
 TIKMSRHDRELKGPLVYVLSITSCITYMSSPALAVICNLGADMDIYER
 RCTREKLPYNNKSPAGSIGMAGTGLASVAVMYFPASFGVIBDSGMILRFVLISA
 SALVESLPISIDIDNLTISLRLAGFLP"
ORIGIN
 Alignment Scores:
 Pred. No.: 4,45e-38 Length: 1360
 Score: 530.50 Matches: 119
 Percent Similarity: 59.0% Conservative: 67
 Best Local Similarity: 37.8% Mismatches: 108
 Query Match: 34.5% Indels: 21
 Gaps: 7
 DB: 15
 US-10-634-548-2 (1-304) x AKI76217 (1-1360)

Db 288 AATATCCGAGCTCTTCACTGCGCCGGTATGTTCCGGAATTCGGTTTATCAGAT 347
 Qy Val-----GlyAlaThrValAlaValIleuGlyGlyValAlaThrAlaValIleuSer 87
 Db 348 GTTCGCGCGCTTTGAGAGTCACTAGCATCGTT-----GCGTTCTGTCGCTTC 398
 Qy PheGluSerLeuThrIlyrArgAenValIleGlnIleSerLeuSerArgIlyrValIleHis 107
 Db 399 TGGGAGAGATGGCAAAACGTGGCATCTTGACACAGAAATCTATCCCGAAGCTTGTCAT 458
 Qy IleLeuSerIlyLeuLeuPheValIleAlaIleAlaIleProIlePheSerGlySerThrGluAla 127
 Db 459 ATAAATATTTGGCTAGATTATTTATGCTTGTGCGCGCTGTTCAGTTCTGGATCCAAAGCA 518
 Qy ArgTyrPheAlaIlePheValIleProLeuValIleAsnGlyLeuArgIleValIleAsnGlyLeu 147
 Db 519 GCACATTTCGCATCTCTTGAACCTGAGCTCAATATAGATAGCATATTCCTGGGCTT 578
 Qy SerIleSerProAsnSerMetLeuIleIlyrSerValThrArgGluIlyrArgAlaGluGlu 167
 Db 579 GAGGTATACACAGACAGAAAGCAATCAATCAATGACATGACAGACATGCGAGAGGAA 638
 Qy LeuLeuIlyGlyProLeuPheTyrValIleAlaIleLeuLeuPheSerAlaValPhePheTyr 187
 Db 639 CTACTTAAAGGCGCGCTTATCTATGATCTGATCAATCAATCAATGACGCTCATCTACTATTTGG 698
 Qy ArgGluSerProIleGlyMetIleSerLeuAlaMetMetCysGlyGlyIlyrAspGlyIleAla 207
 Db 699 AATATATCCCATTCGAGATCGGATGCGATATGCAACTTTGCGAGAGATGATGATGCT 758
 Qy AspIleMetGlyArgIlyrPheGlySerThrIlyrIleProTyrAsnProArgIlyrSerTyr 227
 Db 759 GACATTTGGGTGCGCGCTTTGGAAACAGAAAGCTCTTCAACAAACAAATTCATTT 818
 Qy AlaGlySerIleSerMetPheIlePheGlyPhePheIleSerIleAlaLeuLeuTyrTyr 247
 Db 819 GCTGTAGCATTTGGAATGGACCGCGCGGTTCATGACATCTGTGGATATGATGACTAC 878
 Qy TyrSerSerLeuGlyTyrLeuHisMetAsnTyrPgluThrThrLeuGlnArgValAlaMet 267
 Db 879 TTTCCTCATTTGTTTATCATCGAGATGAGCGGGGAATGATTTCT---CGTTTCCTGTC 935
 Qy ValSerMetValAlaThrValIleGluSerLeuProIleThrAspGlnIleLeuAspAspAsn 287
 Db 936 ATCTCTATACATCAAGCTTGTGGATATCACTCCCAATAGACACCGCATTCAGACACAT 995
 Qy IleSerValProLeuAlaThrIleLeuAlaIleTyrIleuSerPhe 302
 Db 996 CTCACCAATTCCTTAACCTTGCTGCGCGATTCCTTACTCTTC 1040

RESULT 12
 AY087555 1162 bp mRNA linear PLN 14-APR-2003
 DEFINITION Arabidopsis thaliana clone 36602 mRNA, complete sequence..
 ACCESSION AY087555
 VERSION AY087555.1 GI:21406293
 KEYWORDS
 SOURCE
 ORGANISM
 Arabidopsis thaliana (thale cress)
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 1 (bases 1 to 1162)
 Haas, B.J., Volkovskiy, N., Town, C.D., Troukhan, M., Alexandrov, N.,
 Feldmann, K.A., Flavell, R.B., White, O. and Salzberg, S.L.
 Full-length messenger RNA sequences greatly improve genome
 annotation
 Genome Biol. 3 (6), RESEARCH0029 (2002)
 12093376
 2 (bases 1 to 1162)
 Brover, V., Troukhan, M., Alexandrov, N., Lu, Y.-P., Flavell, R. and
 Feldmann, K.
 Full-length cDNA from Arabidopsis thaliana

JOURNAL REFERENCE	JOURNAL
Unpublished (bases 1 to 1162)	
Brover, V., Troukhan, M., Alexandrov, N., Lu, Y.-P., Flavell, R. and Feldmann, K.	
Direct Submission	
Submitted (11-MAR-2002)	Ceres, Inc, 3007 Malibu Canyon Road, Malibu, CA 90265, USA
COMMENT	This clone sequence is one of 5,000 Ceres full-length cDNAs made available to RIGB and Genbank. The following quality assessment of this set was done by comparison with known proteins: two percent of the clones are estimated to be 5'-truncated; less than one percent are 3'-truncated; approximately two percent represent alternative splice variants, including unspliced introns and spliced exons; one percent may contain premature stop codons; five percent may have frame shifts in a coding region. A sequence is considered to be 5'-truncated if it lacks the translation initiation start (ATG). A sequence is considered to be 3'-truncated if it lacks the C-terminal end of the encoded protein. Please note that these cDNA sequences are derived from the Ws or Laer ecotypes and therefore may contain polymorphisms when compared to sequences from Col-0. Genus carried out the library production and sequencing of the full-length clones. Ceres, Inc. carried out the clustering of the 5' sequences, selection of clones, and sequence assembly.
FEATURES	location/Qualifiers
source	1..1162
	/organism="Arabidopsis thaliana"
	/mol_type="mRNA"
	/db_xref="taxon:3702"
	/clone="36602"
	23..946
CDS	/codon_start=1
	/product="unknown"
	/protein_id="AA05097.1"
	/db_xref="GI:215931148"
	/translation="MATTSTTKLVLCSPFSLPVDSPSLAFSPDPFPLTVRIRITSPSSSLPAPTKIRKSLAIVMPENSVLDVAFGVTSLVAPSCIGPMEDIKRGIDPQKLRLRVINIGLVPMLCMPFSSSIQGLRSLVPGANTRMILIGIVYHDEETFGKLSRRHGRRELKGPLYVLVLTSCITVYMYSSPLAIVTCLCGDDMAIVGRTFKMRLPYNNKSPFAGISGMATGFLASVGYMYFASFPIDSGMILRFLVLSIRL SALVESLPISYDIDNLTISLTLAAGLEF"
ORIGIN	
Alignment Scores:	
Pred. No.:	8,32e-38
Score:	526.50
Percent Similarity:	58.7%
Best Local Similarity:	37.5%
Query Match:	34.3%
	Indels: 21
	Gaps: 7
US-10-634-548-2 (1-304) x AY087555 (1-1162)	
QY	1 McaAlaAaThrLeuProLeuSerProIleAenHISgInLeuCyAArgPhedLyAsnAn 20
Db	23 ATGGCACTACTAGTACTACTACCAAGGCTCCGTTCTGCTGCTCTTCAATTCATCT 82
QY	21 SerLeuThrThiAsArgPheCySerPro-----GlyPheLeuIleSerSerPro 37
Db	83 CCTCTCGTTAC-----TCTCTCTCTCTCTCGGCTTCTTCTCCGATTTCA 130
QY	38 CyAPheIleGlyLeu-----ThrGlyMetGlySerAlaThrGInLeuArgAlaArg 54
Db	131 CGATTCCTCATGTCGCAATCGCGACTAGTTAGATGACGCTCAGGCTCCGCGACAC 190
QY	55 ArgSerLeuIleSerSerAlaValAlaThr-----AsnSerLeuLeuHisArg 70
Db	191 AAAATCCGCAAGCTTCACTCGCGCCCGTGATGTTCCGAAAATATCGATTATCAAT 250
QY	71 Val-----GlyAlaThrValAlaValLeuGlyAlaTyrAlaLeuValLeuSer 87
Db	251 GTCTGCGCGGTTTGGAGTCACTAGCATCGTT-----GCGTTCTCGGCTCGGTTTC 301
QY	88 PheGInSerLeuThrThyLArgAsnValIleGInGInSerLeuSerArgLyLeuValHis 107

[illegible]

Yoshino, M. and Hayashizaki, Y.
Collection, mapping, and annotation of over 28,000 cDNA clones from
japonica rice
Science 301 (5631), 376-379 (2003)
12869764
2 (bases 1 to 3143)
2 (bases 1 to 3143)
Fujishiro, T., Aizawa, K., Akimura, T., Arahawa, T., Carninci, P., Doi, K.,
Nishikawa, T., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W.,
Hayashida, K., Hayashizaki, Y., Hayatsu, N., Hiramoto, K., Hirooka, T.,
Hori, F., Hotta, I., Iida, J., Iida, J., Ikeda, R., Imamura, K.,
Imotani, K., Ishibiki, J., Ishii, Y., Ishikawa, M., Itoh, M., Kagawa, I.,
Kawagawa, S., Katoh, H., Kawagashira, N., Kawai, J., Kawamata, M.,
Kikuchi, S., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M.,
Kodama, T., Kohira, K., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
Koya, S., Kurihara, C., Kurosaki, T., Kusumegi, T., Li, C., Lu, M.,
Masuda, H., Matsubara, K., Matsuyama, T., Mura, J., Miyazaki, A.,
Mizuno, K., Murakami, K., Murata, M., Nagata, T., Nakamura, M.,
Naniki, T., Nariawa, R., Nishikawa, J., Nishi, K., Nomura, K.,
Numasaki, R., Ohneda, E., Ohno, M., Ohtsuki, K., Oka, M., Ooka, H.,
Osato, N., Ota, Y., Otsu, Y., Ryu, R., Saitoh, H., Sakai, C., Sakai, K.,
Sakazume, N., Sano, H., Sasaki, D., Sato, K., Sato, K., Shibata, K.,
Shinagawa, A., Shiraki, T., Shishiki, T., Sogabe, Y., Sugano, S.,
Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, M., Tagami-Takeda, Y.,
Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A.,
Toya, T., Tsunoda, Y., Ueda, M., Waki, K., Xie, Q., Yanagi, W.,
Yamada, H., Yamamoto, M., Yasunishi, A., Yazaki, J., Yokomizo, S. and
Yoshimura, A.
Direct Submission
Submitted (05-DEC-2001) Shoshi Kikuchi, National Institute of
Agrobiological Sciences, Department of Molecular Genetics, Head of
Laboratory of Gene Expression, 2-1-2 Kannondai, Tsukuba, Ibaraki
305-8602, Japan (E-mail:skikuchi@nias.affrc.go.jp)
Tel:81-29-838-7007, Fax:81-29-838-7007
This clone is one of the 28K full-length cDNA clones from japonica
rice
URL: http://cdna01.dna.affrc.go.jp/cDNA/
NIHS Rice Full-length cDNA Project Team: Kikuchi, S., Sato, K.,
Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J.,
Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T.,
Ohneda, E., Yanagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T. and
Yamamoto, M.
RAIS Genome Sequencing & Analysis Group: Otsu, Y., Iida, Y.,
Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M., Kobayashi, M.,
Kodama, T., Kurosaki, T., Kusumegi, T., Lu, M., Masuda, H., Mura, J.,
Mizuno, K., Nariawa, R., Nishikawa, J., Oka, M., Ryu, R., Sugano, S.,
Sugiyama, A., Suzuki, Y., Tsunoda, Y., Ueda, M., Xie, Q., Yokomizo, S.,
Yoshimura, A., Matsubara, K. and Murakami, K.
Genome Exploration Research Group in Riken Genomic Sciences Center
and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K.,
Akimura, T., Arahawa, T., Carninci, P., Fukuda, S., Hanagaki, T.,
Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K.,
Hirooka, T., Hori, F., Iida, J., Imamura, K., Imotani, K., Ishii, Y.,
Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawai, J.,
Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Mura, M.,
Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Osato, N.,
Ota, Y., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H.,
Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T.,
Sogabe, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F.,
Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Waki, K.,
Yasunishi, A. and Hayashizaki, Y.
Location/Qualifiers
1. .3143
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cui="F013046L12"
/db_xref="taxon:39947"
/clone="J013046L12"

ORIGIN
Alignment Scores: 3.03e-37 Length: 3143
Score: 526.50 Matches: 126

Percent Similarity: 50.1% Conservative: 43
Best Local Similarity: 37.4% Mismatches: 37
Query Match: 34.3% Indels: 133
DB: 15 Gaps: 4
US-10-634-548-2 (1-304) x AK065924 (1-3143)
QY 99 GlnSerLeuSerArgIysLeuValHisIleLeuSerGlyLeuLeuPheValIleAlaTyr 118
Dp 2920 CAGAACCTGAGCAGAAAATTGTCATGTCTCTCCGAGTCTCTGTCATGTCTCTG 2861
QY 119 ProlIePhe----- 121
Dp 2860 CCACTATTCAGGTCCTCCCTTCCACATGCTCTTCCACATGCAAAATTCTTCAACCA 2801
QY 122 -----SerGlySerThrGluAlaAr 128
Dp 2800 AATCAGACTGAGATCTTAAGAAAATCATGCCCTTCCGAGTAAATTCAGACAGACAG 2741
QY 128 GTYrPheAlaIlePheValProlLeuValAangIysLeuValIleangIysLeu 148
Dp 2740 GTTCTTCGCGCAATTGTCCGTTGCTGAATGCAATGAGCTTCTGACCTACGCGCTCCG 2681
QY 148 rIleSerProAnsSerMetLeuIleYsSerValThrArgGluGlyArgAlaGlu----- 166
Dp 2680 CTTTCCACTGATGAGAGCTCTAGTAAATCGTGAACCGTGAAGAAAACCAAGATTCT 2621
QY 166 ----- 166
Dp 2620 GCTATGATCTCTTATTTCTGATCTCTTCAATCAAAAATAGATATCTTGCAT 2561
QY 167 -----GluLeuLeuYsGlyProlLeuPheTyrValLeuAlaLeu 180
Dp 2560 GATCTTGTGTTGCTTACAGGAAATGCTGAGAGGCTCTGTAATGATCATGTGCTGC 2501
QY 180 eupheSerAlaValPhePheTyrArgIysSerProlIleGlyMetIleSerLeuAlaMet 200
Dp 2500 TGGTCACTGTTTGTGCTTCTTGGGCGACATCTCTATGAGATTTGTTGTTGTCATGA 2441
QY 200 ecYsGlyGlyAaPglYIle----- 206
Dp 2440 TGAAGTGTGATGATGATGA-GATGTTGATCTCTCGGAATTAGAGAAATTCGACATT 2382
QY 206 ----- 206
Dp 2381 TTTTGTATATACCCAGTTCATGATGAAATGATGCTTCTGCTAGTATATCCACT 2322
QY 207 -----AlaSerIleMetGlyArgGlyPheGlySerThr 218
Dp 2321 TCTTGTCTTATTTTTCAGGCTTGTCTCAATGTTGGAGAGATATGCTCCGCA 2262
QY 218 yslleProTyrAsnProArgIysSerTrpAlaGlySerIleSerMetPheIlePheGlyP 238
Dp 2261 AGCTGCATTCATGAGAAACAGAGCTGATAGAGAACATCTCAATTCTTCTGCT 2202
QY 238 hePheIleSerIleAla----- 243
Dp 2201 TCCTTTATCTGAC-ACGTAGAGTTTTCCTCAGATGATGATGCTGCTTCTTTCAT 2143
QY 244 -----LeuLeuTyrTyrTyr 249
Dp 2142 TTAATCTTCTTATTTGTAATCTAGATCTTTCATCTACAGAGATGCTGTTACTTCT 2083
QY 249 erSerLeuGlyTyrLeuHisMetAsnTrpGluThrThrLeuGluArgValAlaMetValS 269
Dp 2082 CTTCGCTTGTGTTACTTACATGTCGTGCGATGCGACATTCGTAATCTGCTTGTG 2023
QY 269 erMetValAlaThrValValGluSerLeuProlIleThrArgGluLeuAspAspAsnIleS 289
Dp 2022 CATTAAGCACTACTAGTGAAGTATGATCTCTGTCATGATGTTGTGATGACAAATATCT 1963
QY 289 erValProlLeuAlaThrIleLeuAlaAlaTyrLeuSerPheGlyTyr 304

Dd	100016	TACCGGGGACGTTCCGG---	ATCTCGTGGCGGTGGCGGTGGAAATCGCTCCCGGTC	99960
Qy	281	ThzrpgpGlnuepApApApAnlleSerVal	ProleuAlaThrllleuAlaAlaTryleu	300
Dd	99959	TCATCCGCTTTGGACGACAACTTCA	CGCTGTCTTACGCGCGCTGGCGTTGGGACGTTG	99900
Qy	301	Serpe	302	
Dd	99899	TTATTTC	99894	

Search completed: March 16, 2006, 22:17:27
Job time : 7927 secs

THIS PAGE LEFT BLANK

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: March 16, 2006, 19:53:40 ; Search time 681 Seconds

(without alignments) 2975.134 Million cell updates/sec

Title: US-10-634-548-2

Perfect score: 1536

Sequence: 1 MATULPSPINHQLCRGNN.....DDNISVPLATILAAVLSFGY 304

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4996997 seqs, 333346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Command line parameters:

-MODE=frame+ .p2n.model -DEV=xjh
-Q=/abs/ABSSWEB.spool/US10634548/runat.16032006.084431.10295/app_query.fasta.1
-DB=N.GeneSeq -QFMT=fastcap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000 -HOST=abs02n
-USRR=US10634548 -OCGN 1.1 727 @runat.16032006.084431.10295 -NCPU=6 -ICPU=3
-NO_MMAP -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-MAIN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FCGAPOP=6 -FCGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_GeneSeq_21.*

1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*
14: geneseqn2005as:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1536	100.0	1015	13	ADx46288
2	1536	100.0	1091	12	ADJ98163
3	1523	99.2	980	13	ADx31786
4	1523	99.2	1088	3	AAc36634

5	1284.5	83.6	1897	12	ADJ98165	AdJ98165 Thale cre
6	1266	82.4	1062	13	ADx54389	Adx54389 plant ful
7	870	56.6	1007	13	ADx45670	Adx45670 plant ful
8	794	51.7	1085	13	ADx47586	Adx47586 plant ful
9	788	51.3	1033	13	ADx51853	Adx51853 plant ful
10	776	50.5	798	13	ADx48515	Adx48515 plant ful
11	773	50.3	976	13	ADx51811	Adx51811 plant ful
12	751	48.9	723	13	ADx50931	Adx50931 plant ful
13	751	48.9	753	13	ADx51758	Adx51758 plant ful
14	742	48.3	937	13	ADx51714	Adx51714 plant ful
15	648	42.2	494	3	AAc53042	AAc53042 Zea maye
16	646	42.1	674	13	ADx34057	Adx34057 plant ful
17	573	37.3	916	13	ADx17675	Adx17675 plant ful
18	572.5	37.3	1225	13	ADx51722	Adx51722 plant ful
19	567.5	36.9	688	13	ADx47647	Adx47647 plant ful
20	566.5	36.9	666	13	ADx10350	Adx10350 plant ful
21	552.5	36.0	1091	13	ADx51558	Adx51558 plant ful
22	545.5	35.5	1241	13	ADx19540	Adx19540 plant ful
23	540.5	35.2	795	13	ADx30592	Adx30592 plant ful
24	539	35.1	867	13	ADx50633	Adx50633 plant ful
25	536.5	34.9	1046	13	ADx36475	Adx36475 plant ful
26	530.5	34.5	1160	3	AAc45477	AAc45477 Arabidops
27	530.5	34.5	1162	12	ADJ98167	AdJ98167 Thale cre
28	517	33.7	1193	13	ADx15971	Adx15971 plant ful
29	515	33.5	963	3	AAc46147	AAc46147 Arabidops
30	512	33.3	1029	13	ADx32079	Adx32079 plant ful
31	511.5	33.3	648	13	ADx34109	Adx34109 plant ful
32	509	33.1	944	13	ADx37159	Adx37159 plant ful
33	509	33.1	1106	13	ADx54040	Adx54040 plant ful
34	508	33.1	948	13	ADx37160	Adx37160 plant ful
35	492	32.0	1165	3	AAc40207	AAc40207 Arabidops
36	490	31.9	1044	13	ADx51812	Adx51812 plant ful
37	481.5	31.3	699	13	ADx50711	Adx50711 plant ful
38	479.5	31.2	793	13	ADx47651	Adx47651 plant ful
39	466	30.3	545	13	ADx45744	Adx45744 plant ful
40	440	28.6	1507	13	ADx34329	Adx34329 plant ful
41	430.5	28.0	1502	13	ADx45999	Adx45999 plant ful
42	392.5	25.6	776	13	ADx63685	Adx63685 plant ful
43	381.5	24.8	729	13	ADx60081	Adx60081 plant ful
44	366.5	23.9	631	13	ADx49148	Adx49148 plant ful
45	353	23.0	485	5	AAh87900	AAh87900 Peppermin

ALIGNMENTS

RESULT 1	ADx46288	standard; cDNA; 1015 BP.
ID	ADx46288	
XX	ADx46288	
AC	ADx46288	
XX		
DT	21-APR-2005	(first entry)
XX		
DE	Plant full length insert polynucleotide seqid 21028.	
XX		
KW	plant protectant; plant growth regulant; gene therapy; plant;	
KW	recombinant DNA construct; physical array; plant breeding marker;	
KW	cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;	
KW	extreme osmotic condition; pathogen tolerance; pest tolerance;	
KW	growth rate; cell cycle pathway; disease resistance;	
KW	galactomannan production; lignin production; plant growth regulator;	
KW	yield; plant growth; plant development; seed oil; protein yield;	
XX	protein content; gene; ss.	
XX		
OS	Unidentified.	
XX		
PN	US2004034888-A1.	
XX		
PD	19-FEB-2004.	
XX		
PF	28-APR-2003; 2003US-00425114.	
XX		
PR	06-MAY-1999; 99US-00304517.	

PR 05-NOV-2001; 2001US-00985678.
 XX (Liu/J) Liu J.
 PA (ZHOU/) ZHOU Y.
 PA (KOVA/) KOVALIC D K.
 PA (SCRE/) SCREEN S E.
 PA (TABAS/) TABASKA J E.
 PA (CAOV/) CAO Y.
 XX
 PI Liu J, Kovalic DK, Screen SE, Tabaska JE, Cao Y;
 DR WPI; 2004-180133/17.
 XX
 PT New recombinant DNA construct, useful for improving plant tolerance to
 PT cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or
 PT pests, for conferring increased resistance to plant disease, or for
 PT improving yield.
 XX
 PS Claim 1; SEQ ID NO 21028; 15pp; English.
 XX
 CC The invention describes a recombinant DNA construct comprising a
 CC polynucleotide consisting of a sequence encoding an amino acid sequence
 CC available in electronic form from the US patent office at
 CC ftp.segdata.uspto.gov/sequence.html?docID:200403488. The polynucleotide
 CC of the invention are also useful in physical arrays of molecules and as
 CC plant breeding markers. The recombinant DNA construct is useful for
 CC improving plant tolerance to cold, heat, drought, herbicides, extreme
 CC osmotic conditions, pathogens or pests, for manipulating growth rate in
 CC plant cells by modification of the cell cycle pathway, for conferring
 CC increased resistance to plant disease, for producing galactomannan,
 CC lignin or plant growth regulators, for increasing the rate of homologous
 CC recombination in plants, for improving yield by modification of
 CC photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake
 CC or by providing improved plant growth and development under at least one
 CC stress condition or for modifying seed oil or protein yield and/or
 CC content. This sequence represents a plant full length insert
 CC polynucleotide that can be used in the recombinant DNA construct of the
 CC invention.
 XX
 SQ Sequence 1015 BP; 228 A; 245 C; 219 G; 323 T; 0 U; 0 Other;
 XX
 Alignment Scores:
 Pred. No.: 9,79e-166 Length: 1015
 Score: 1536.00 Matches: 304
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 13 Gaps: 0
 US-10-634-548-2 (1-304) x ADX46288 (1-1015)
 QY 1 MetTAAATThrLeuProLeuSerProTLeaAnhIsgInLeuCYARGPhgLYAAsn 20
 DB 35 ATGGGAGGAACTTACCTTATCTCCGATCAATCTCGTGTGCTGGGAAACAC 94
 QY 21 SerLeuThrThrnhIARGPhgCYSerProGLYPhLeuIleSerSerProCYpHeIle 40
 DB 95 TCTTGAAGACTCAACGGCTTCTGTTCTCTGGCTTCTGATTCTTCTCTGTTTATT 154
 QY 41 GlyLeuThrGLYMetGLYSerAlaThrGLInLeuArgIlaARGaGSerLeuIleSerSer 60
 DB 155 GGTGTGACCGGAATGGCTGTCTACTACTAGTTACCTGCTCGTCTCTATCTCTTCA 214
 QY 61 AlaValAlaThrAnSerLeuLeuHISArValAGIYAlaThrValAlaValLeuGLY 80
 DB 215 GCAGTGGGAGCAAAATTCCTGTGCATACGTCGAGGCAACCGTGGCACTGTGGTGA 274
 QY 81 AlaTyrAlaLeuValIleuSerPhgIuSerLeuThrLYSARGAsnValIleGLInSer 100
 DB 275 GCATACGGCTTGTCTTAACTTGAAGTCTCACCAAGCGAAAGTCAATTCAACAGAGT 334
 QY 101 LeuSerArgLYSLeuValHISLleuSerGLYLeuLeuPheValIleuAlaTrpProIle 120

DB 335 TTGAGCAGAAAGCTGTGTCATATATCTTCAGGCTGCTTTTCGTACTGCGTGGCAATC 394
 QY 121 PheSerGLYSerThrGluAlaARGTYrPheAlaAlaPheValProLeuValAsnGLYLeu 140
 DB 395 TTCAAGCGGATCGAACCGAGGCTTCATCTTGTCTGCTTTTGTTCGTTAGCAAGGCTTA 454
 QY 141 ArgLeuValIleAnGLYLeuSerIleSerProAnSerMetLeuIleLYSArValThr 160
 DB 455 AGGCTGTATTAAACGAGATATCATTTCCCAAAATTCGATGATCAATCAATCCGTACA 514
 QY 161 ArgGLYGLYArgAlaGLYLeuLeuLeuLYSGLYProLeuPheTYrValIleuAlaLeuLeu 180
 DB 515 AGAGAAAGGAGAGAGAGAGAGTGTCTTAAAGTCTTGTGTCTTACGTTCTTACGTTCTT 574
 QY 181 PheSerAlaValPhePheTYrARGGLYSerProIleGLYMetIleSerLeuAlaMetMet 200
 DB 575 TTCTCTGCGGTTTCTTCTTGAAGAGACTCTCATGCGTATGATCTGTTAGCATATAGT 634
 QY 201 CYaGLYGLYAspGLYIleAlaAspIleMetGLYArgLYSPheGLYSerThrLYSLeuPro 220
 DB 635 TGTGTGGCGATGGAGAAATGCTGATATATATGAGACGTAAAGTTGGGTCAACTAAGATACCT 694
 QY 221 TYrAsnProARGLYSerTYrAlaGLYSerIleSerMetPheIlePheGLYPhePheIle 240
 DB 695 TACAACCCAAAGAAAGAGTGTGGCGAGAGATCTCCATGTCATCTTCGCTTCTTCAATC 754
 QY 241 SerIleAlaLeuLeuTYrTYrTYrSerSerLeuGLYTYrLeuHISMetAsnTrpGLYThr 260
 DB 755 TCCATGCCATTAATCTTACTTACTTACTCAAGCTTGGGATACCTTCAACATGAGAAACG 814
 QY 261 ThrLeuGLNArgValAlaMetValSerMetValAlaThrValAlaGLYSerLeuProIle 280
 DB 815 ACCTTGAGAGAGAGAGCAATGTCTCAATGTGTCCGACGAGTACGAGTCCGTACCATC 874
 QY 281 ThrAspGLNLeuAspAspAsnIleSerValProLeuAlaThrIleLeuAlaAlaTYrLeu 300
 DB 875 ACCGATTAATTAGCGCAAAATTTCCGTTCTGCTGCTACTATTTTAACTCTTATTTA 934
 QY 301 SerPheGLYTYr 304
 DB 935 AGTTTCGATAT 946
 RESULT 2
 ADJ98163
 ID ADJ98163 standard; DNA; 1091 BP.
 XX
 AC ADJ98163;
 XX
 DT 06-MAY-2004 (first entry)
 XX
 DE Thale cress LTT1 phycol kinase wild-type DNA.
 XX
 KW phycol kinase; tocopherol biosynthesis; plant; drought resistance; LTT1;
 KW thale cress; ds; gene; wild-type.
 XX
 OS Arabidopsis thaliana.
 XX
 FN MO2004013312-A2.
 XX
 PD 12-FEB-2004.
 XX
 PE 05-AUG-2003; 2003WO-US025276.
 XX
 PR 05-AUG-2002; 2002US-0400689P.
 PR 05-AUG-2003; 2003US-00634548.
 XX
 PA (MONS) MONSANTO TECHNOLOGY LLC.
 XX
 PI Norris SR, Lincoln K, Abad MS, Ellere R, Harteuylker KK;
 PI Hirschberg J, Karunanandaa B, Moshiri F, Stein JC, Valentin HE;
 PI Venkatesh TV;
 XX
 DR WPI; 2004-157125/15.

DR P-PSDB; ADJ98164.
XX New phyto. kinase polynucleotides, useful in mediating tocopherol
PT biosynthesis and in producing plants with increased drought resistance.
XX
XX Claim 4; SEQ ID NO 1; 189pp; English.
XX
CC The invention relates to a novel substantially purified nucleic acid
CC molecule encoding a phyto. kinase. The nucleic acid molecules and
CC polypeptides of the invention may be useful in mediating tocopherol
CC biosynthesis and in producing plants with increased drought resistance.
CC The current sequence is that of the chate cross LRT1 phyto. kinase wild-
CC type DNA of the invention.
XX
XX Sequence 1091 BP; 252 A; 260 C; 223 G; 356 T; 0 U; 0 Other;
SQ
Alignment Scores:
Pred. No.: 1,09e-165 Length: 1091
Score: 1536.00 Matches: 304
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 12 Gaps: 0
US-10-634-548-2 (1-304) x ADJ98163 (1-1091)
QY 1 MetAlaIaThrLeuProLeuSerProIleAsnHISGlnLeuCyArgPheGlyAsnAsn 20
DB 108 ATGGACACCACTTACCTTATCTCCGATCATATCATCGTGTGCGGTGGGAAACAC 167
QY 21 SerLeuThrThrHISArgPheCySerProGlyPheLeuIleSerSerProCyPheIle 40
DB 168 TCTTTGACGACTCACCGGTTGTCTTCTCCGCTTCTTGAATTTCTTCTCTGTTTCA 227
QY 41 GlyLeuThrGlyMetGlySerAlaThrGlnLeuArgAlaArgSerLeuIleSerSer 60
DB 228 GGTGACCGGAAAGGCTGTGCTACTGAGTACGTCGTCGTCGTCGTCGTCGTCGTC 287
QY 61 AlaValAlaThrAsnSerLeuLeuHISAspValGlyAlaThrValAlaValLeuGly 80
DB 288 GCAATTCGACGAAATTCGCTGTCGATGACGTCGACGACGACGTCGTCGTCGTCG 347
QY 81 AlaTyrAlaLeuValLeuSerPheGlySerLeuThrIleAsnValIleGlnGlnSer 100
DB 348 GCATACCGGCTGTCTTATGCTTCAAGATGTCACCAAGCAACGTCATTCACAGAG 407
QY 101 LeuSerArgIleLeuValHISIleLeuSerGlyLeuLeuPheValLeuAlaTrpProIle 120
DB 408 TTGAGCAGAAAGCTTGATGATATCTCTCAGGTCGCTTTCGTAATTCGTCGTC 467
QY 121 PheSerGlySerThrGlnAlaArgTyrPheAlaAlaPheValProlLeuValAsnGlyLeu 140
DB 468 TTCAGCGAATCGACGAGGCTCGATCTTGTCTTGTTCGTTAGTAAGGCTTGA 527
QY 141 ArgLeuValIleAsnGlyLeuSerIleSerProAsnSerMetLeuIleLeuSerValThr 160
DB 528 AGCGTTGTTATTAACGACATCTCCATTTCCCAATTCGATGCTAAATCAATCCGTC 587
QY 161 ArgGlnIleArgAlaGlnLeuLeuLeuValGlyProlLeuPheValLeuAlaLeuLeu 180
DB 588 AGAAGAGGAG 647
QY 181 PheSerAlaValAlaPhePheTrpArgIleLeuSerProIleGlyMetIleSerLeuAlaMetMet 200
DB 648 TTCCTCTCGGTTTCTTCTGAGAGAGCTCTCTTCCGTAAGATCTCGTTAGCAATATG 707
QY 201 CyseGlyIleAspGlyIleAlaAspIleMetGlyArgIlePheGlySerThrIlePro 220
DB 708 TGCTGTGCGCATGAGAAATGCTGATTAATGAGAGAGAGAGAGAGAGAGAGAGAGAG 767
QY 221 TyrAsnProArgIleSerTrpAlaGlySerIleSerMetPheIlePheGlyPhePheIle 240
DB 768 TACAACCAAGAAAGAGTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 827

QY 241 SerIleAlaLeuLeuTyrTyrTyrSerSerLeuGlyTyrLeuHISMetAsnTrpGluThr 260
DB 828 TCCATCGATTAATCTTATCTTATCTTCAAGCTTGGGATCTTACAGATGAGACTGGGAAACG 887
QY 261 ThrLeuGlnArgValAlaMetValSerMetValAlaThrValValGlySerLeuProIle 280
DB 888 ACCTTGACAG 947
QY 281 ThrAspGlnLeuAspAspAsnIleSerValProlLeuAlaThrIleLeuAlaAlaTyrLeu 300
DB 948 ACCATCAATTAAGACGACAAATATTCGTTCTCTGCTACTATTAATTAATTAATTA 1007
QY 301 SerPheGlyTyr 304
DB 1008 AGTTGCGATAT 1019
RESULT 3
ADX31786
ID ADX31786 standard; cDNA; 980 BP.
XX
XX ADX31786;
XX
XX 21-APR-2005 (first entry)
XX
XX
XX Plant full length insert polynucleotide seqid 14606.
XX
XX plant protectant; plant growth regulator; gene therapy; plant;
XX recombinant DNA construct; physical array; plant breeding marker;
XX cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;
XX extreme osmotic condition; pathogen tolerance; pest tolerance;
XX growth rate; cell cycle pathway; disease resistance;
XX galactomannan production; lignin production; plant growth regulator;
XX yield; plant growth; plant development; seed oil; protein yield;
XX protein content; gene; ss.
XX
XX Unidentified.
XX
XX US2004034888-A1.
XX
XX 19-FEB-2004.
XX
XX 28-APR-2003; 2003US-00425114.
XX
XX 06-MAY-1999; 99US-00304517.
XX
XX 05-NOV-2001; 2001US-00985678.
XX
XX (LIU/) LIU J.
XX (ZHOU/) ZHOU Y.
XX (KOVA/) KOVALIC D K.
XX (SCRE/) SCREEN S E.
XX (TABA/) TABASKA J E.
XX (CAOY/) CAO Y.
XX
XX
XX Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JE, Cao Y;
XX WPI; 2004-180133/17.
XX
XX New recombinant DNA construct, useful for improving plant tolerance to
XX cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or
XX pests, for conferring increased resistance to plant disease, or for
XX improving yield.
XX
XX Claim 1; SEQ ID NO 14606; 15pp; English.
XX
XX The invention describes a recombinant DNA construct comprising a
XX polynucleotide consisting of a sequence encoding an amino acid sequence
XX available in electronic form from the US patent office at
XX ftp.segdata.uspto.gov/sequence.html?docid=2004034888. The polynucleotide
XX of the invention are also useful in physical arrays of molecules and as
XX plant breeding markers. The recombinant DNA construct is useful for
XX improving plant tolerance to cold, heat, drought, herbicides, extreme
XX osmotic conditions, pathogens or pests, for manipulating growth rate in

CC plant cells by modification of the cell cycle pathway, for conferring
 CC increased resistance to plant disease, for producing galactomannan,
 CC lignin or plant growth regulators, for increasing the rate of homologous
 CC recombination in plants, for improving yield by modification of
 CC photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake
 CC or by providing improved plant growth and development under at least one
 CC stress condition or for modifying seed oil or protein length and/or
 CC content. This sequence represents a plant full length insert
 CC polynucleotide that can be used in the recombinant DNA construct of the
 CC invention.

XX SQ Sequence 980 BP; 223 A; 229 C; 215 G; 313 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	2,87e-164	Length:	980
Score:	1523.00	Matches:	301
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	99.2%	Indels:	0
DB:	13	Gaps:	0

US-10-634-548-2 (1-304) x ADX31786 (1-980)

Qy 4 ThrLeuProLeuSerProIleAsnHisGlnLeuCyArgPheGlyAaMaSLeuThr 23
 Db 3 ACCTTACCTTATCTCCGATCATCATCAGTTGTGCGTTGGGAAACAACCTTTGACG 62
 Qy 24 ThrHisArgPheCySerProGlyPheLeuIleSerProCyPheIleGlyLeuThr 43
 Db 63 ACTCACCGGTTCTTCTCCGCGCTTCTGATTTCTTCTTTCATATGGTTTACCC 122
 Qy 44 GlyMetGlySerAlaThrGlnLeuArgAlaArgArgSerLeuIleSerSerAlaValAla 63
 Db 123 GGAATGGGCTTCTGCTACTCACTTACGTCGTCGTTCTCGACTTTCAGCGAGTTGCG 182
 Qy 64 ThrAnSerLeuLeuHisAspValGlyAlaThrValAlaValLeuGlyValAlaTyrAla 83
 Db 183 ACGAATCGCTGTGATGATGATCGAGCCACCTGGCAGCTGGTGGAGCATACGCG 242
 Qy 84 LeuValLeuSerPheGlySerLeuThrIlyArgAnValIleGlnGlnSerLeuSerArg 103
 Db 243 CTTGTCTTAAGCTTGCAGAGTCTCACCAAGCAACGATTCATCAAGAGTTTGAAGCA 302
 Qy 104 LysLeuValHisIleLeuSerGlyLeuLeuPheValLeuAlaTyrProIlePheSerGly 123
 Db 303 AAGCTTGGCATATCTCAAGTCTGCTTTTCGACTTGGTGGCCATCTTCAAGCGGA 362
 Qy 124 SerThrGluAlaArgTyrPheAlaAlaPheValProLeuValAaGlyLeuArgLeuVal 143
 Db 363 TCGACCGAGGCTCGATACCTTGTGCTTGTTCGTTAGTGAATGCGCTTAAGGCTTGT 422
 Qy 144 IleAnGlyLeuSerIleSerProAnSerMetLeuIleLeuSerValThrArgGluGly 163
 Db 423 ATTAAACGACATCATTCATTTCCCAATTCGATGCTAATCAATCCGTCACAAGAAAGG 482
 Qy 164 ArgAlaGluGlyLeuLeuValGlyPheLeuPheTyrValLeuAlaLeuLeuPheSerAla 183
 Db 483 AGACCAAGAAAGTTCCTAAAGTCTTGTCTTCAAGTTCAGCTCTTCTTCTGCGG 542
 Qy 184 ValPhePheTyrArgGlySerProIleGlyMetIleSerLeuAlaMetMetCysGlyGly 203
 Db 543 GTTTCTCTCGAGAGAGTCTCTATCGGTATGATCTCGTTAGCAATATGATGTGTGCG 602
 Qy 204 AspGlyIleAlaAspIleMetGlyArgGlyPheGlySerThrIlyValIleProTyrAspPro 223
 Db 603 GATGAAATAGCTGATATATGAGAGTATGATTTGGGTCACTAAGATACCTTACACCCA 662
 Qy 224 ArgLysSerTyrAlaGlySerIleSerMetPheIlePheGlyPhePheIleSerIleAla 243
 Db 663 AGAAAGATTTGGCGAGAGCATCTCCATGTTCACTTCGCTTTCATCTCCATCCCA 722
 Qy 244 LeuLeuTyrTyrTyrSerSerLeuGlyTyrLeuHisMetAsnTyrGluThrThrLeuGln 263

Db 723 TTACTTACTTATTACTCAAGCTTGGGTACTTTCACATGAATCGGAAACGACCTTGACG 782
 Qy 264 ArgValAlaMetValSerMetValAlaThrValValGlySerLeuProIleThrAspGln 283
 Db 783 AGGTGCAATAGTCTTCATATGTCGCGACGATGTCAGTGCATCCCATCACCGATCA 842
 Qy 284 LeuAspAspAniIleSerValProLeuAlaThrIleLeuAlaAlaTyrLeuSerPheGly 303
 Db 843 TTAGACGACAAATTTTCGTTCTGCTGCTAGCTACTATTACTGCTTATTAAGTTTCGA 902
 Qy 304 Tyr 304
 Db 903 TAT 905
 RESULT 4
 AAC36634
 ID AAC36634 standard; DNA; 1088 BP.
 XX AAC36634;
 AC 17-OCT-2000 (first entry)
 XX
 DE Arabidopsis thaliana DNA fragment SEQ ID NO: 14519.
 XX
 KW Hybridisation assay; genetic mapping; gene expression control;
 KW protein identification; signal transduction pathway; metabolic pathway;
 KW promoter; termination sequence; ss.
 XX
 OS Arabidopsis thaliana.
 XX
 PN EP1033405-A2.
 PD 06-SEP-2000.
 XX
 PF 25-FEB-2000; 2000EP-00301439.
 XX
 XX 25-FEB-1999; 99US-0121825P.
 PR 05-MAR-1999; 99US-0123180P.
 PR 09-MAR-1999; 99US-0123548P.
 PR 23-MAR-1999; 99US-0125788P.
 PR 25-MAR-1999; 99US-0126264P.
 PR 29-MAR-1999; 99US-0126785P.
 PR 01-APR-1999; 99US-0127462P.
 PR 08-APR-1999; 99US-0128234P.
 PR 16-APR-1999; 99US-0128714P.
 PR 19-APR-1999; 99US-0129845P.
 PR 21-APR-1999; 99US-0130077P.
 PR 23-APR-1999; 99US-0130449P.
 PR 28-APR-1999; 99US-0130510P.
 PR 30-APR-1999; 99US-0130891P.
 PR 04-MAY-1999; 99US-0131449P.
 PR 05-MAY-1999; 99US-0132048P.
 PR 06-MAY-1999; 99US-0132485P.
 PR 07-MAY-1999; 99US-0132486P.
 PR 11-MAY-1999; 99US-0132487P.
 PR 14-MAY-1999; 99US-0132487P.
 PR 15-MAY-1999; 99US-0132487P.
 PR 18-MAY-1999; 99US-0132487P.
 PR 19-MAY-1999; 99US-0132487P.
 PR 20-MAY-1999; 99US-0132487P.
 PR 21-MAY-1999; 99US-0132487P.
 PR 24-MAY-1999; 99US-0132487P.
 PR 25-MAY-1999; 99US-0132487P.
 PR 27-MAY-1999; 99US-0132487P.
 PR 28-MAY-1999; 99US-0132487P.
 PR 01-JUN-1999; 99US-0132487P.
 PR 03-JUN-1999; 99US-0132487P.

PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142054P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0143277P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145090P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 05-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.

PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148119P.
PR 12-AUG-1999; 99US-0148314P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150844P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151067P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 13-OCT-1999; 99US-0159296P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159332P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160771P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Alignment Scores:
Pred. No.:

3,36e-164

Length:

1088

Score: 1523.00 Matches: 300
 Percent Similarity: 99.7% Conservative: 3
 Best Local Similarity: 98.7% Mismatches: 1
 Query Match: 99.2% Indels: 0
 DB: 3 Gaps: 0

US-10-634-548-2 (1-304) x AAC36634 (1-1088)

QY 1 MetAlaAlaThrLeuProLeuSerProIleAsnHISGlnLeuCyArgPheGlyAsnAsn 20
 DB 108 ATGGCAGACCACTTAATCTATCCGATCAATCATAGTGTGTGGTTCGGGAACAC 167
 QY 21 SerLeuThrThrHisArgPheCysSerProGlyPheLeuIleSerSerProCysPheIle 40
 DB 168 TCTTGAACGACTCACCGGTTCTGTCTCGGCTTCTGATTCTTCTCTCTGTTTCACTT 227
 QY 41 GlyLeuThrGlyMetGlySerAlaThrGlnLeuArgAlaArgGlySerLeuIleSerSer 60
 DB 228 GGTTCACCGGAATGGGCTGTGCTACTCATGTTACGTCGTCGTCCTCATCTCTTCA 287
 QY 61 AlaValAlaThrAsnSerLeuLeuHisAspValGlyAlaThrValAlaValLeuGlyGly 80
 DB 288 GCAGTTGGACGAATTCCTCTGTGCATACGTCGAGCCACCGTGGCAGTCTTGTGGA 347
 QY 81 AlaTyralaLeuValLeuSerPheGluSerLeuThrHisArgAsnValIleGlnGlnSer 100
 DB 348 GCATACCGGCTTGTCTTAAGTTCGAGAGTCTCACCGAAGCAACGTCATTCAACAGGT 407
 QY 101 LeuSerArgGlyLeuValHisIleLeuSerGlyLeuLeuPheValLeuAlaTrpProIle 120
 DB 408 TTGAGCAGAAAGCTGTGTCATATCTCAGGTCTGCTTTCGTACTTGCCTGGCCATC 467
 QY 121 PheSerGlySerThrGlnAlaArgTyrrPheAlaIlePheValProLeuValAsnGlyLeu 140
 DB 468 TTCACGGATGACCGAGGCTCGATCTTGTGCTTGTTCCTTGTGTGTAAGGCTTA 527
 QY 141 ArgLeuValIleAsnGlyLeuSerIleSerProAsnSerMetLeuIleLeuSerValThr 160
 DB 528 AGGCTGTGTTATTTAACGACTATCCATTTCCCAATTCGATGCTATCAATCCGTCACA 587
 QY 161 ArgGlnGlyArgAlaGlnGlnLeuLeuLeuGlyProLeuPheTyrrValLeuAlaLeuLeu 180
 DB 588 AGAAGAGGAGAGCAGAAAGTGTGTTAAGTCTTGTGTTCTTCGTTCTGAGCTTCTT 647
 QY 181 PheSerAlaValPhePheTrpArgGluSerProIleGlyMetIleSerLeuAlaMetMet 200
 DB 648 TTCTCTGCGGTTTCTTCTGAGAGAGTCTCTATCGATAGATCTCGTTAGCAATGATG 707
 QY 201 CysGlyGlyAspGlyIleAlaAspIleMetGlyArgGlyPheGlySerThrIleLeuPro 220
 DB 708 TGTGGTGGCGATGGAAATGCTGATATTAATGGGACGTAAAGTTGGGTCAACTAAGTACT 767
 QY 221 TyrAsnProArgGlySerTrpAlaGlySerIleSerMetPheIlePheGlyPhePheIle 240
 DB 768 TACAAACCCAGAAAGAGTTGGGACAGAGCAATCTCAAGTTCATCTTGGCTTCTTCA 827
 QY 241 SerIleAlaLeuLeuTyrrTyrrSerSerLeuGlyTyrrLeuHisMetAsnTrpGlnThr 260
 DB 828 TCCATCGCATTAATCTTACTACTACAGCTTGGGTACCTTCACTTCACTTAACGGAAACG 887
 QY 261 ThrLeuGlnArgValAlaMetValSerMetValAlaThrValAlaGluSerLeuProIle 280
 DB 888 ACCTTGACGAGAGTGAATAGTCTCAATGTGACCAAGTGTGAGTGGCTACCAACATC 947
 QY 281 ThrAspGlnLeuAspAspAsnIleSerValProLeuAlaThrIleLeuAlaAlaTyrrLeu 300
 DB 948 ACCGATCAATTAAGACAGCAACGTTTCGTTCTCTGCGTACTATTATTAACTGCTTATTTA 1007
 QY 301 SerPheGlyTyrr 304
 DB 1008 AGTTTCGATAT 1019

RESULT 5

ADJ98165
 ID ADJ98165 standard; DNA; 1897 BP.
 AC ADJ98165;
 XX

06-MAY-2004 (first entry)

Thale cress LTT1 phycol kinase mutant DNA.

XX phycol kinase; tocopherol biosynthesis; plant; drought resistance; LTT1;
 KM thale cress; ds; gene; mutant.

XX Arabidopsis thaliana.

XX MO200401312-A2.

XX 12-FEB-2004.

XX 05-AUG-2003; 2003MO-US025276.

XX 05-AUG-2002; 2002US-040689P.

XX 05-AUG-2003; 2003US-00634548.

XX (MONS) MONSANTO TECHNOLOGY LLC.

XX Norris SR, Lincoln K, Abad MS, Eilers R, Hartsuyker KK;
 PI Hirschberg J, Karunanandaa B, Moshiri F, Stein JC, Valentin HE;
 PI Venkatesh TV;

XX WPI; 2004-157125/15.

XX P-PSDB; ADJ98166.

XX New phycol kinase polynucleotides, useful in mediating tocopherol
 PT biosynthesis and in producing plants with increased drought resistance.

XX Example 2; SEQ ID NO 3; 189pp; English.

XX The invention relates to a novel substantially purified nucleic acid
 CC molecule encoding a phycol kinase. The nucleic acid molecules and
 CC polypeptides of the invention may be useful in mediating tocopherol
 CC biosynthesis and in producing plants with increased drought resistance.

CC The current sequence is that of the thale cress LTT1 phycol kinase mutant
 CC DNA of the invention.

XX SQ Sequence 1897 BP; 501 A; 408 C; 332 G; 656 T; 0 U; 0 Other;

Alignment Scores: 1.7e-136 Length: 1897
 Pred. No.: 1284.50 Matches: 301

Score: 1284.50 Conservative: 0

Percent Similarity: 52.4% Mismatches: 3

Best Local Similarity: 52.4% Indels: 272

Query Match: 12 Gaps: 5

DB: US-10-634-548-2 (1-304) x ADJ98165 (1-1897)

QY 1 MetAlaAlaThrLeuProLeuSerProIleAsnHISGlnLeuCyArgPheGlyAsnAsn 20

DB 108 ATGGCAGACCACTTAATCTATCCGATCAATCATAGTGTGTGGTTCGGGAACAC 167

QY 21 SerLeuThrThrHisArgPheCysSerProGlyPheLeuIleSerSerProCysPheIle 40

DB 168 TCTTGAACGACTCACCGGTTCTGTCTCGGCTTCTGATTCTTCTCTGTTTCACTT 227

QY 41 GlyLeuThrGlyMetGlySerAlaThrGlnLeuArgAlaArgGlySerLeuIleSerSer 60

DB 228 GGTTCACCGGAATGGGCTGTGCTACTCATGTTAGTGTGCTGTTCTGTGATCTTCA 287

QY 61 AlaValAlaThrAsnSerLeuLeuHisAspValGlyAlaThrValAlaValLeuGlyGly 80

DB 288 GCAGTTGGACGAATTCCTCTGTGCATACGTCGAGCCACCGTGGCAGTCTTGTGGA 347

QY 81 AlaTyralaLeuValLeuSerPheGluSerLeuThrHisArgAsnValIleGln----- 98

D	b		 	348	GCAATCGGGCTTGTCTTAAGTTCCAGAGTCCTCACACGGAAGCTCATTTACAACAGTGC	407
O	y	98	- - - - -	98	- - - - -	98
D	b		 	408	TCTTAATAATCGTTTTAGTTATCCACAAATTTCTCCGTTTACAATTCCAGTTTATTGC	467
O	y	98	- - - - -	98	- - - - -	98
D	b		 	468	AACACCACTAATATGTTGAAGAAGATTTCTCAGTGTGTTTGCAAGTACTCATTTAGAA	527
O	y	98	- - - - -	98	- - - - -	98
D	b		 	528	ACAATGATTAAGCTTAGAAATTTTGTGTGAATTAGTTTTCATTTCGAAATTTTATTA	587
O	y	98	- - - - -	98	- - - - -	98
D	b		 	588	GAATTTGTACACCTTAGTAAGCAATATPCACTTATCATGCCAATCGTAAGACCGA	647
O	y	98	- - - - -	98	- - - - -	98
D	b		 	648	CAGAACAAGATGTCCAAAAATATTTACCGCTTATATGTTACCACCTTTTCTAACCTC	707
O	y	98	- - - - -	98	- - - - -	98
D	b		 	708	CCTTTAACTATCCGTAATCGCTTACCGCTAAAAATATACCGTTCCTTGTGTAAACA	767
O	y	99	- - - - -	99	- - - - -	99
D	b		 	768	AAGTAAGAAAAGAAAGAAACAATACTTGATGTGTTTATGTGTAGACAGAGTTTGAACG	827
O	y	103	GLyLeuValHisIleLeuSerGlyLeuLeuPheValLeuAlaTrpProIle-	120	- - - - -	120
D	b		 	828	AAAGCTTGCAATATACCTCTCAGGCTGTGCTTTCCGTACTGCGTGGCCAATCTTCAGGTA	887
O	y	120	- - - - -	120	- - - - -	120
D	b		 	888	TTGCTTCTCTCTATGTTGTGAATCTCTGTACCTTTTAAACATGATAGCATTTCTGA	947
O	y	121	- - - - -	121	- - - - -	121
D	b		 	948	TTTCTTTTACTCATCTTAAGTTATAGCGGATCGACCGAGCTCATATCTTGTGCTTT	1007
O	y	133	eValProLeuValIleangLyLeuArgLeuValIleangLyLeuSerLieserProAnse	153	- - - - -	153
D	b		 	1008	TGTTCCGTTAGTAAGCTTAAGGCTTGTTATTAACGCACTATCCATTTCCCCAAATTC	1067
O	y	153	rMetLeuIleLysSerValThrArgGluGlyArgAlaGlu-	166	- - - - -	166
D	b		 	1068	GATGCTAATCAAATCCGTCACAAAGAAAGGAGCAGAGTAAGTGTCTAAGTTTTTTT	1127
O	y	166	- - - - -	166	- - - - -	166
D	b		 	1128	TCCAACCTTGATATGATTTTTCACAATCTGATTCACACTTTCTTGTTTTCCAACATCA	1187
O	y	167	- - - - -	167	- - - - -	167
D	b		 	1188	CAGAGAGTTGGCTTAAGGCTCTTGTTGTCTAGCTTCTAGCTCTTCTTCTCGGGTTTT	1247
O	y	185	ePheTrpArgGluSerProIleGlyMetLieserLeuAlaMetMetCyrgLyGlyAspGI	205	- - - - -	205
D	b		 	1248	CTTCGGAGAGAGTCTCCTATCGGTATATCTCGTTAGCAATGATGTGTGGCGAATGG	1307
O	y	205	YIle-	206	- - - - -	206
D	b		 	1308	TAA-ATTTCTGTCAAGTACTAAGTATTAATTAATTAACAATTTTACAAATGCCATPA	1367
O	y	207	- - - - -	207	- - - - -	207
D	b		 	1367	ATGTAATAACTAAGTGCATCAATATGTCTATGTAGAAATGCTGATATATATGGAGCG	1426
O	y	212	GLyPheGlySerThrLygIleProTyraAsnProArgLysSerTrpAlaGlySerLise	232	- - - - -	232

Db	1427	TAAGTTGGGCGCACTAAGACTTACCTTCAACCCAGAAAGAGTTAGGACAGAAACATCTC	1486
Qy	232	rmcPheIIlePheGlyPhePheIIleSerIIleAla-----	243
Db	1487	CATGTTCACTTCGCGCTTCTTCATCTCCATCGC-GTAAATAATATCAATCCACTATT	1545
Qy	243	-----	243
Db	1546	AATCATCAAAATGTCCTCTCTTGTGACGACGACAAAGCTTATAGAACTGAGATGAGTTGC	1605
Qy	244	-----LeuIleuTYrTYrSerSerIle	251
Db	1606	TACTTAACTCAACCGTTTCTTCTTTGTAAATTTTGGAGATTACTTACTTATTACTCAACCT	1665
Qy	251	uGIYrIleuHisMetAnTpgIurThrIleuGlnArgValAlaMetValSerMetVa	271
Db	1666	TGGGTAACCTTCACTGAAGTGGAAAGCACTTSCAAGAGTAGTACATGCTCAATGGT	1725
Qy	271	IAIaThrValValGlnSerIleuProlIetrArgGlnLeuAspAspSerIleSerValPr	291
Db	1726	CGCACAAGTACTGAGTCGCTACCCATCACCAGATCAATTGAGACACAATATTTCGGTTCC	1785
Qy	291	oleuAlaThrIleleuAlaAlaTYrLeuSerPheGlyTYr	304
Db	1786	TCTGGCTACTATTATTAGCTGCTATTATTAAAGTTTCGGATAT	1825
RESULT 6			
ID	ADX54389		
AC	ADX54389 standard; cDNA; 1062 BP.		
XX	ADX54389;		
DT	21-APR-2005 (first entry)		
XX			
DE	Plant full length insert polynucleotide seqid 29129.		
KM	plant protectant; plant growth regulant; gene therapy; plant;		
KM	recombinant DNA construct; physical array; plant breeding marker;		
KM	cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;		
KM	extreme osmotic condition; pathogen tolerance; pest tolerance;		
KM	growth rate; cell cycle pathway; disease resistance;		
KM	galactomannan production; lignin production; plant growth regulator;		
KM	yield; plant growth; plant development; seed oil; protein yield;		
KM	protein content; gene; ss.		
OS	Unidentified.		
XX			
PN	US2004034888-A1.		
PD	19-FEB-2004.		
XX			
PF	28-APR-2003; 2003US-00425114.		
XX			
PR	06-MAY-1999; 99US-00304517.		
PR	05-NOV-2001; 2001US-00985678.		
XX			
PA	(LIU//) LIU J.		
PA	(ZHOU//) ZHOU Y.		
PA	(KOVA//) KOVALIC D K.		
PA	(SCRE//) SCREEN S E.		
PA	(TABAA//) TABASKA J E.		
PA	(CAOY//) CAO Y.		
XX			
PI	Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JE, Cao Y;		
XX	WPI; 2004-180133/17.		
XX			
PT	New recombinant DNA construct, useful for improving plant tolerance to		
PT	cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or		
PT	improving yield.		
XX			
PS	Claim 1; SEQ ID NO 29129; 15pp; English.		

XX The invention describes a recombinant DNA construct comprising a
 CC polynucleotide consisting of a sequence encoding an amino acid sequence
 CC available in electronic form from the US patent office at
 CC ftp.secdat.uspto.gov/sequence.html?docID:2004034888. The polynucleotide
 CC of the invention are also useful in physical arrays of molecules and as
 CC plant breeding markers. The recombinant DNA construct is useful for
 CC improving plant tolerance to cold, heat, drought, herbicides, extreme
 CC osmotic conditions, pathogens or pests, for manipulating growth rate in
 CC plant cells by modification of the cell cycle pathway, for conferring
 CC increased resistance to plant disease, for producing galactomannan,
 CC lignin or plant growth regulators, for increasing the rate of homologous
 CC recombination in plants, for improving yield by modification of
 CC photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake
 CC or by providing improved plant growth and development under at least one
 CC stress condition or for modifying seed oil or protein yield and/or
 CC content. This sequence represents a plant full length insert
 CC polynucleotide that can be used in the recombinant DNA construct of the
 CC invention.

XX Sequence 1062 BP; 235 A; 264 C; 243 G; 320 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	Score:	Length:	Matches:
9.39e-135	1266.00	1062	250
Percent Similarity:	89.3%	Conservative:	24
Best Local Similarity:	81.4%	Mismatches:	27
Query Match:	82.4%	Indels:	6
DB:	13	Gaps:	2

US-10-634-548-2 (1-304) x ADX54389 (1-1062)

QY 1 MetAlaIaThrLeuProLeuSerProIleAsnHisGlnLeuCyArgPheGlyAsnAsn 20
 DB 46 ATGGCGGACGCTTATCTTATCTCCGGTACCATCGATTATGCGATAAGCAACAGG 105
 QY 21 -----SerLeuThrHisArgPheCySerProGlyPheLeuIleSerSerPro 37
 DB 106 TTCTGGTAAACGATATGATCCCGGTTCTGTTGCCA-----GTTTCTTCGCCC 156
 QY 38 CysPheIleGlyLeuThrGlyMetGlySerAlaThrGlnLeuAlaArgArgSerLeu 57
 DB 157 TGTTCATCGCGGATAAGAAATCGGCTCGCATCGATTACGCGGCTCGCACCCCTTG 216
 QY 58 ILeSerSerAlaValAlaThrAsnSerLeuHisAspValGlyAlaThrValAlaVal 77
 DB 217 ATCTTTCAGGCGCTTCATCATATCTATTCATGACGTCGAGCCACGCTGCACTT 276
 QY 78 LeuGlyGlyAlaTyraIleuValIleuSerPheGlnSerLeuThrLysArgAsnValIle 97
 DB 277 CTTAGTGGTGCTTACGCGCTTGTCTTACTCTTCAGAGCTTCACGAGACGATGATT 336
 QY 98 GlnGlnSerLeuSerArgLysLeuValHisIleLeuSerGlyLeuLeuPheValLeuAla 117
 DB 337 CCACAGAGATTGAGAGAAAGCTTGTGATATATCTCAGGCTCTTTCGCGCTTCG 396
 QY 118 TrpProIlePheSerGlySerThrGlnAlaArgTyraPheAlaAlaPheValProLeuVal 137
 DB 397 TGGCCAACTTCACGCGCATCAACGAGGCTCGATCTTTCGCTTTGCTTCCTTAAG 456
 QY 138 AsnGlyLeuArgLeuValIleAsnGlyLeuSerIleSerProAsnSerMetLeuIleLys 157
 DB 457 AATGGCCTTAAGGCTTGTCTCAACGTTTGTCTCCCTCCCTAACCTCCACGCTATCCA 516
 QY 158 SerValThrArgGlnGlyArgAlaGlnGlnLeuLysGlyProLeuPheTyraValLeu 177
 DB 517 TCCGTAACCTAGGAGAGAGACCAAGAGTTGCTTAAAGGTCATTGTTCTACCTTCA 576
 QY 178 AlaLeuLeuPheSerAlaValPhePheTrpArgGlnSerProIleGlyMetIleSerLeu 197
 DB 577 GCCCTTCAGTTCGTCGCACTTTCTTCGAGAGATTCTCCCTACCGGATATATACGCTG 636
 QY 198 AlaMetMetCysGlyGlyAspGlyIleAlaAspIleMetGlyArgLysPheGlySerThr 217

DB 637 GCATATGTTGTGGTGGCGACGGAATGCTGATATCATGGACGTACGATCATATAC 696
 QY 218 LysIleProTyraAsnProAspGlySerTrpAlaGlySerIleSerMetPheIlePheGly 237
 DB 697 AAGATACCTTACCAACCCAGAAAGAGCTTGGCGGAGACATCTTCATCTCATTTTGGC 756
 QY 238 PhePheIleSerIleAlaLeuLeuTyraTyraTyraSerSerLeuGlyTyraLeuHisMetAsn 257
 DB 757 TTCTTCATTCCTCAACCGATTACTTACTTATTAATTAATCAAGCTTACGGTATCTTCACATGAA 816
 QY 258 TrpGlnThrThrLeuGlnAlaValAlaMetValSerMetValAlaThrValAlaGlnSer 277
 DB 817 TGGGAAACCAACCTTTACAGAGAGTCGCGATTGTCTCATTTGCTGCTACGTTGGTAGCTCA 876
 QY 278 LeuProIleThrAspGlnLeuAspAspAsnIleSerValProLeuAlaThrIleLeuAla 297
 DB 877 CTACCCATCAACCGATCAAAATAGACGACAGCTTCGATTCTCTGGCTACTATTTTGGCT 936
 QY 298 AlaTyraLeuSerPheGlyTyra 304
 DB 937 GCTTACCTTAAGTTTGGATAT 957

RESULT 7

ADX45670
 ID ADX45670 standard; cDNA; 1007 BP.

AC ADX45670;

DT 21-APR-2005 (first entry)

XX Plant full length insert polynucleotide seqid 20410.

DE plant protectant; plant growth regulant; gene therapy; plant;
 KW recombinant DNA construct; physical array; plant breeding marker;
 KW cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;
 KW extreme osmotic condition; pathogen tolerance; pest tolerance;
 KW growth rate; cell cycle pathway; disease resistance;
 KW galactomannan production; lignin production; plant growth regulator;
 KW yield; plant growth; plant development; seed oil; protein yield;
 KW protein content; gene; ss.

OS Unidentified.

XX US2004034888-A1.

XX 19-FEB-2004.

XX 28-APR-2003; 2003US-00425114.

XX 06-MAY-1999; 99US-00304517.

XX 05-NOV-2001; 2001US-00985678.

XX (LNU/) LNU J.

XX (ZHOV/) ZHOV Y.

XX (KOVA/) KOVALIC D K.

XX (SCRE/) SCREEN S E.

XX (TABA/) TABASKA J E.

XX (CAOV/) CAO Y.

XX Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JE, Cao Y;

XX WPI; 2004-180133/17.

XX New recombinant DNA construct, useful for improving plant tolerance to

XX cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or

XX pests, for conferring increased resistance to plant disease, or for

XX improving yield.

XX Claim 1; SEQ ID NO 20410; 15pp; English.

XX The invention describes a recombinant DNA construct comprising a

CC polynucleotide consisting of a sequence encoding an amino acid sequence

CC plant cells by modification of the cell cycle pathway, for conferring
 CC increased resistance to plant disease, for producing galactomannan,
 CC lignin or plant growth regulators, for increasing the rate of homologous
 CC recombination in plants, for improving yield by modification of
 CC photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake
 CC or by providing improved plant growth and development under at least one
 CC stress condition or for modifying seed oil or protein yield and/or
 CC content. This sequence represents a plant full length insert
 CC polynucleotide that can be used in the recombinant DNA construct of the
 CC invention.

CC Sequence 1033 BP, 170 A; 287 C; 294 T; 0 U; 0 Other;

CC Alignment Scores:

Pred. No.:	Score:	Length:	Matches:	Mismatches:	Indels:	Gaps:
Best Similarity:	788.00	1033	160	60	76	13
Best Local Similarity:	51.84	Conservative:	60	76	13	13
Query Match:	51.34	Mismatches:	60	76	13	13
DB:	13	Indels:	13	76	13	13
		Gaps:	3	76	13	13

US-10-634-548-2 (1-304) x ADX51853 (1-1033)

QY 6 ProLeuSerProIleAsnHleuGlnLeuCyArg-----PheGlyAsnAsnSerLeu 22
 DB 6 CCCGTCGACGTGCTCGGACCTTCCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 65
 QY 23 -----ThrThriAArgPheCySerProGlyPheLeuIleSerSer-ProCy 38
 DB 66 CTGCTGTCGCGCTCCAGTCAGGCTCGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 125
 QY 38 ePheIleGlyLeu--ThrGlyMetGlySerAlaThrGlnLeuAlaArgArgSerLeu 57
 DB 126 CGGCGAAGGCTGTGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGG 176
 QY 58 IleSerSerAlaValAlaThrAsnSerLeuHleuAspValAlaIleThrValAlaVal 77
 DB 177 GCGCTCGGCGGCTCGGCGGCTCGGCGGCTCGGCGGCTCGGCGGCTCGGCGGCTCGG 236
 QY 78 LeuGlyGlyAlaTyraLeuValLeuSerPheGlnSerLeuThrIleArgAsnValIle 97
 DB 237 ACCGCGGCGGCTGCT 296
 QY 98 GlnGlnSerLeuSerArgGlyLeuValHleIleLeuSerGlyLeuLeuPheValLeuAla 117
 DB 297 GAACGAACTGAGGAAAGAAATTTGATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 356
 QY 118 TrpProIlePheSerGlySerThrGlnAlaArgTyraPheAlaAlaPheValProLeuVal 137
 DB 357 TGGCCACTATTCAATTCAGACAGACAGACAGGCTCTCTCTCTCTCTCTCTCTCTCTCT 416
 QY 138 AsnGlyLeuArgLeuValIleAsnGlyLeuSerIleSerProAsnSerMetLeuIleYs 157
 DB 417 AACTGCATTAAGGCTTCTGACCTACGAGGCTCCGCTTCCATGATGAGAGCTCTAGTAAA 476
 QY 158 SerValThrArgGlnGlyArgAlaGlnLeuLeuValGlyProLeuPheTyraValLeu 177
 DB 477 TCGGAGACCCGAGGAAAGAAACAGAGAAATTTGAGAGGCTCTCTCTCTCTCTCTCTCT 536
 QY 178 AlaLeuLeuPheSerAlaValPhePheTrpArgIleSerProIleGlyMetIleSerLeu 197
 DB 537 GTGCTGCTGCTGCTGCTTGTGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 596
 QY 198 AlaMetMetCyGlyGlyAspGlyIleAlaAspIleMetGlyArgValPheGlySerThr 217
 DB 597 TCGATGATGAGTGTGCTGATGCTTGTGCTGATGCTTGTGCTGATGCTTGTGCTGCTG 656
 QY 218 LysIleProTyraAsnProArgIleSerTrpAlaGlySerIleSerMetPheIlePheGly 237
 DB 657 AAGCTGCCATTAAGAGAAAGAGAGCTGATGATGATGATGATGATGATGATGATGATGAT 716
 QY 238 PhePheIleSerIleAlaLeuLeuTyraTyraSerSerLeuGlyTyraLeuHleuMetAsn 257

DB 717 TTCCTTATATCGACATGATGCTGTCTTACTTCTCTGCTTGTACTTCACTCTGCG 776
 QY 238 TrpGlnThrThrLeuGlnArgValAlaMetValSerMetValAlaThrValAlaIleSer 277
 DB 777 TGGGATCTGCGACCTTGTGAAACCTGCTCTTGTGATTAAGACAGCTACGTATGAGTGT 836
 QY 278 LeuProIleThrAspGlnLeuAspAsnIleSerValProLeuAlaThrIleLeuAla 297
 DB 837 ATTCTGTCAATGATGTTGTGATGATCAATATCTCCGTTCTTGTGCGCACATGTTGGCA 896
 QY 238 AlaTyLeuSerPheGlyTyr 304
 DB 897 GCTATCTGTATTATTGGCTAC 917
 RESULT 10
 ADX48515
 ID ADX48515 standard; cDNA, 798 BP.
 XX
 AC ADX48515;
 XX
 DT 21-APR-2005 (first entry)
 XX
 DB Plant full length insert polynucleotide seqid 23255.
 XX
 KW plant protectant; plant growth regulant; gene therapy; plant;
 KW recombinant DNA construct; physical array; plant breeding marker;
 KW cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;
 KW extreme osmotic condition; pathogen tolerance; pest tolerance;
 KW growth rate; cell cycle pathway; disease resistance;
 KW galactomannan production; lignin production; plant growth regulator;
 KW yield; plant growth; plant development; seed oil; protein yield;
 KW protein content; gene; ss.
 XX
 OS unidentified.
 XX
 PN US2004034888-A1.
 XX
 PD 19-FEB-2004.
 XX
 PF 28-APR-2003; 2003US-00425114.
 XX
 PR 06-MAY-1999; 99US-00304517.
 XX
 PR 05-NOV-2001; 2001US-00985678.
 XX
 PA (LNUJ)/ LNU J.
 PA (ZHOU)/ ZHOU Y.
 PA (KOVA)/ KOVALIC D K.
 PA (SCRE)/ SCREEN S E.
 PA (TABAKA)/ TABASKA J E.
 PA (CAOY)/ CAO Y.
 XX
 PI Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JE, Cao Y;
 XX
 DR WPI, 2004-180133/17.
 XX
 PT New recombinant DNA construct, useful for improving plant tolerance to
 PT cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or
 PT pests, for conferring increased resistance to plant disease, or for
 PT improving yield.
 XX
 PS Claim 1; SEQ ID NO 23255; 15bp; English.
 XX
 CC The invention describes a recombinant DNA construct comprising a
 CC polynucleotide consisting of a sequence encoding an amino acid sequence
 CC available in electronic form from the US patent office at
 CC ftp://seqdata.uspto.gov/sequence.html?docid:2004034888. The polynucleotide
 CC of the invention are also useful in physical arrays of molecules and as
 CC plant breeding markers. The recombinant DNA construct is useful for
 CC improving plant tolerance to cold, heat, drought, herbicides, extreme
 CC osmotic conditions, pathogens or pests, for manipulating growth rate in
 CC plant cells by modification of the cell cycle pathway, for conferring
 CC increased resistance to plant disease, for producing galactomannan,
 CC lignin or plant growth regulators, for increasing the rate of homologous


```

Query Match:      50.3%      Indels:      18
DB:               13        Gaps:         1
US-10-634-548-2 (1-304) x ADX51811 (1-976)

QY      33 LeuIIeSerSerProCyAphelIeGlyLeuThrGlyMetGlySerAlaThrGlnLeuArg 52
DB      60 CTCCTCCTCCCTCCGACCTCGCTGCTGCGCGCTGCACCTCCGCCCCGACCTCGCC 119
QY      53 AlaArgArg----- 55
DB      120 GCCCGAGGCGCCAGGCGGTGCTGCTCGCGCGCGCGCTCCGCGCTGCGCGGCGCG 179
QY      56 ---SerLeuIIeSerSerAlaValAlaThrAnsSerLeuLeuHisArgValAlaThr 74
DB      180 CTGGGCGCTCGGCTCGACGCGCGCGCGCTCGCTCGCTCGCTCGGACGCGCGCGCGCACG 239
QY      75 ValAlaValLeuGlyValAlaThrAlaLeuValLeuSerPheGlnSerLeuThrArg 94
DB      240 CTGCTCGTACCGCGCGCGCTCATCTCCCTCGCGCGCTTGACGCGCTCACGCGCGCC 299
QY      95 AenValIIeGlnGlnSerLeuSerArgGlyLeuValHisIleLeuSerGlyLeuLeuPhe 114
DB      300 CGCCTCGTCCACAGAGCTTGAGCAGAGAGGTGCTGCTATCCGCGGCTCTTTTTC 359
QY      115 ValLeuAlaThrProIlePheSerGlySerThrGlnAlaArgTyrPheAlaIaIaPheVal 134
DB      360 ATGGCTTCAAGGCGCTTTCAGCAATCTTACAGTCAAGTCAAGTCAAGTCAAGTCAAGT 419
QY      135 ProLeuValAenGlyLeuArgLeuValIIeAenGlyLeuSerIleSerProAnsSerMet 154
DB      420 CCGTTCCTGGAATCGGTGAAGCTTCTCACTCAAGGCTCGGCTCTCACTCCGACGACGAGT 479
QY      155 LeuIIeIySerSerValThrArgGlnGlyArgAlaGlnGlnLeuLeuIyGlyProLeuPhe 174
DB      480 CTAGTAAATCGGTGACCGCTGACGAGAAACGAGAGGAAATTCGTCGAGGCGCTCTGTAC 539
QY      175 TyrValIIeAlaLeuLeuPheSerAlaValPhePheThrArgGlnSerProIleGlyMet 194
DB      540 TATGTCATCGTCTACTATGATCACTGTTCTAGTCTTTGGCGGAGCTCCCGCATCGAGATC 599
QY      195 IleSerLeuAlaMetMetCyseGlyGlyArgGlyIleAlaAerPleMetGlyArgIyPhe 214
DB      600 GTTTCCTTGTCATGATGAGCGGTGACGCGCTTTCGACATTCGTCGAGAGAGGTTTC 659
QY      215 GlySerThrIyIleProTyrAnsProArgIySerTTPAlaGlySerIleSerMetPhe 234
DB      660 GCGTCTGCTCAAGCTCCATTCAACAAAGAGAGAGTGGGTGGGAGTGCAGATGTTTC 719
QY      225 IlePheGlyPhePheIleSerIleAlaLeuLeuTyrTyrIySerSerLeuGlyTyrLeu 254
DB      720 ATCTGTGGTCTCTCTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 779
QY      255 HisMetAenTyrGlnThrThrLeuGlnArgValAlaMetValSerMetValAlaThrVal 274
DB      780 CATGTCACCTGGGATCAGCGCGCTTGATACGCTTCTGCTGCTGCTGCTGCTGCTGCTG 839
QY      275 ValGlnSerLeuProIleThrArgGlnLeuAerPheAnsIleSerValProLeuAlaThr 294
DB      840 GTGAGGTATATCTGTAACTGATGTGTAGATGACATATCTGTTCCCTTGGCCACCC 899
QY      295 IleLeuAlaAlaTyrLeuSerPheGly 303
DB      900 ATGTGTGAGCTTCTGTGTTTGGC 926

```

```

DE      Plant full length insert polynucleotide seqid 25671.
XX      KM plant protectant; plant growth regulant; gene therapy; plant;
XX      KM recombinant DNA construct; physical array; plant breeding marker;
XX      KM cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;
XX      KM extreme osmotic condition; pathogen tolerance; pest tolerance;
XX      KM growth rate; cell cycle pathway; disease resistance;
XX      KM galactomannan production; lignin production; plant growth regulator;
XX      KM yield; plant growth; plant development; seed oil; protein yield;
XX      KM protein content; gene; ss.
XX      OS Unidentified.
XX      PN US2004034888-A1.
XX      PD 19-FEB-2004.
XX      PF 28-APR-2003; 2003US-00425114.
XX      PR 06-MAY-1999; 99US-00304517.
XX      PR 05-NOV-2001; 2001US-00985678.
XX      PA (LITUJ/) LIU J.
XX      PA (ZHOU/) ZHOU Y.
XX      PA (KOVA/) KOVALIC D K.
XX      PA (SCRE/) SCREEN S B.
XX      PA (TABAR/) TABASKA J E.
XX      PA (CAOY/) CAO Y.
XX      PI Liu J, Zhou Y, Kovalic DK, Screen SB, Tabaska JB, Cao Y;
XX      DR WPI; 2004-180133/17.
XX      PT New recombinant DNA construct, useful for improving plant tolerance to
XX      PT cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or
XX      PT pests, for conferring increased resistance to plant disease, or for
XX      PT improving yield.
XX      PS Claim 1; SEQ ID NO 25671; 15bp; English.
XX      CC The invention describes a recombinant DNA construct comprising a
XX      CC polynucleotide consisting of a sequence encoding an amino acid sequence
XX      CC available in electronic form from the US patent office at
XX      CC ftp://seeddata.uspto.gov/sequence.html?docid:2004034888. The polynucleotide
XX      CC of the invention are also useful in physical arrays of molecules and as
XX      CC plant breeding markers. The recombinant DNA construct is useful for
XX      CC improving plant tolerance to cold, heat, drought, herbicides, extreme
XX      CC osmotic conditions, pathogens or pests, for manipulating growth rate in
XX      CC plant cells by modification of the cell cycle pathway, for conferring
XX      CC increased resistance to plant disease, for producing galactomannan,
XX      CC lignin or plant growth regulators, for increasing the rate of homologous
XX      CC recombination in plants, for improving yield by modification of
XX      CC photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake
XX      CC or by providing improved plant growth and development under at least one
XX      CC stress condition or for modifying seed oil or protein yield and/or
XX      CC content. This sequence represents a plant full length insert
XX      CC polynucleotide that can be used in the recombinant DNA construct of the
XX      CC invention.
XX      SQ Sequence 723 BP; 145 A; 174 C; 198 G; 206 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:      5.78e-76      Length:      723
Score:          751.00      Matches:      137
Percent Similarity: 83.0%      Conservative: 48
Best Local Similarity: 61.4%      Mismatches:   38
Query Match:    48.9%      Indels:       0
DB:             13        Gaps:          0

US-10-634-548-2 (1-304) x ADX50931 (1-723)
QY      80 GlyAlaTyrAlaLeuValLeuSerPheGlnSerLeuThrIyArgAenValIIeGlnGln 99

```

```

Db      7  GGGCGCTACTCTCTTGTGGCGGCTTTGCAGACGAGTCAACCGAGCGGCGCTCATCGAAAG 66
Qy      100 SerLeuSerArgIleValHisIleLeuSerGlyLeuLeuPheValLeuAlaIlePro 119
Db      67  AGCTTGACGAGAGAGTTGTGACGTCGTCGGCGCTCTGTTTCATGTCATCTTGGGCC 126
Qy      120  IlePheSerGlySerThrGluAlaArgIlePheAlaIlePheValProLeuValAsnGly 139
Db      127  CTGTTACAGCAATTCACACAGACGAGTATTTCCGCCGCGTGTCCCGTCTCTGAACCTCC 186
Qy      140  LeuArgLeuValIleAsnGlyLeuSerIleSerProAsnSerMetLeuIleLysSerVal 159
Db      187  ATGAGCGCTTCTGATATGATGACTCCGCTCTACACTGATGAAGCTCTGTAAATCAAGTG 246
Qy      160  ThrArgGluGlyArgAlaGluGluLeuLeuGlySerGlyProLeuPheThrValLeuAlaLeu 179
Db      247  ACACGTGAAGGAAACCAAGAGAAATTCGAGAGAGTCCACTTATTAATGTCCTGGTGGC 306
Qy      180  LeuPheSerAlaValPhePheTrpArgIleSerProIleGlyMetIleSerLeuAlaMet 199
Db      307  CTGTTACAGCGTTTATGCTCTTGGCGTGAAGTCCCATCGGAGATGCTCTCTGTGCATG 366
Qy      200  MetCysGlyGlyAspGlyIleAlaAspIleMetGlyArgIlePheGlySerThrIle 219
Db      367  ATGAGCGGTGGCGATGTTCCGCTGACATTTGGGAGAGAGTATGGCTCAGCGAAGCTG 426
Qy      220  ProTrpAsnProArgIleSerTrpAlaGlySerIleSerMetPheIlePheGlyPhePhe 239
Db      427  CCATTCACATCGAAGAGAGCGGCGGCGGACATCTCATATTCATTTCTGGTTCTG 486
Qy      240  IleSerIleAlaLeuLeuTrpIleTrpIleSerSerLeuGlyIleLeuHisMetAsnTrpIle 259
Db      487  CTGTCCGCAATGATATGCTCTACTTCTCAACCTCGGTTATGATGATGATGATGCGGAA 546
Qy      260  ThrTrpLeuGlnArgValAlaMetValSerMetValAlaThrValValGluSerLeuPro 279
Db      547  GAGGCACTGTGAAGCGCGCTTGTGTCATGACAGCAAGTGAAGTGAAGCGCTCTCT 606
Qy      280  IleTrpAspGlnLeuAspAsnIleSerValProLeuAlaThrIleLeuAlaIleTrp 299
Db      607  GTGACCGAAGTTGTAGATGACAAATATGTTCTTGGCCACATCGCTGATGCTTTT 666
Qy      300  LeuSerPhe 302
Db      667  CTCTGTTT 675

RESULT 13
ID      ADX51758
XX      ADX51758 standard; cDNA; 753 BP.
AC      ADX51758;
XX      21-APR-2005 (first entry)
DT      Plant full length insert polynucleotide seqid 26498.
XX      DE
XX      XX      Plant full length insert polynucleotide seqid 26498.
KW      plant protectant; plant growth regulator; gene therapy; plant;
KW      recombinant DNA construct; physical array; plant breeding marker;
KW      cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;
KW      extreme osmotic condition; pathogen tolerance; pest tolerance;
KW      growth rate; cell cycle pathway; disease resistance;
KW      galactomannan production; lignin production; plant growth regulator;
KW      yield; plant growth; plant development; seed oil; protein yield;
KW      protein content; gene; sb.
XX      Unidentified.
XX      OS
XX      PN      US2004034888-A1.
XX      PD      19-FEB-2004.
XX      PF      28-APR-2003; 2003US-00425114.
XX

```

```

PR      06-MAY-1999; 99US-00304517.
PR      05-NOV-2001; 2001US-00985678.
XX      PA
XX      PA      (LIU/J) LIU J.
XX      PA      (ZHOU/Y) ZHOU Y.
XX      PA      (KOVA/) KOVALIC D K.
XX      PA      (SCRE/) SCREEN S E.
XX      PA      (TABAS/) TABASKA J E.
XX      PA      (CAO/Y) CAO Y.
PI      Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JE, Cao Y;
DR      WPI; 2004-180133/17.
XX      XX      New recombinant DNA construct, useful for improving plant tolerance to
PT      cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or
PT      pests, for conferring increased resistance to plant disease, or for
PT      improving yield.
XX      PS      Claim 1; SEQ ID NO 26498; 15bp; English.
XX      CC      The invention describes a recombinant DNA construct comprising a
CC      polynucleotide consisting of a sequence encoding an amino acid sequence
CC      available in electronic form from the US patent office at
CC      ftp://seqdata.uspto.gov/sequence.html?docid:2004034888. The polynucleotide
CC      of the invention are also useful in physical arrays of molecules and as
CC      plant breeding markers. The recombinant DNA construct is useful for
CC      improving plant tolerance to cold, heat, drought, herbicides, extreme
CC      osmotic conditions, pathogens or pests, for manipulating growth rate in
CC      plant cells by modification of the cell cycle pathway, for conferring
CC      increased resistance to plant disease, for producing galactomannan,
CC      lignin or plant growth regulators, for increasing the rate of homologous
CC      recombination in plants, for improving yield by modification of
CC      photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake
CC      or by providing improved plant growth and development under at least one
CC      stress condition or for modifying seed oil or protein yield and/or
CC      content. This sequence represents a plant full length insert
CC      polynucleotide that can be used in the recombinant DNA construct of the
CC      invention.
SQ      Sequence 753 BP; 154 A; 180 C; 206 G; 213 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 6,14e-76 Length: 753
Score: 751.00 Matches: 137
Percent Similarity: 83.0% Conservative: 48
Best Local Similarity: 61.4% Mismatches: 38
Query Match: 48.9% Indels: 0
DB: 13 Gaps: 0

US-10-634-548-2 (1-304) x ADX51758 (1-753)
Qy      80  GlyAlaTrpAlaLeuValLeuSerPheGluSerLeuThrIleArgAsnValIleGln 99
Db      16  GGGCGCTACTCTTGTGGCGGCTTTGCAGACGCTCAACGAGCGGCGCTCATCGAAAG 75
Qy      100  SerLeuSerArgIleValHisIleLeuSerGlyLeuLeuPheValLeuAlaIlePro 119
Db      76  AGCTTGACGAGAGAGTTGTGACGTCGTCGGCGCTCTGTTTCATGTCATCTTGGGCC 135
Qy      120  IlePheSerGlySerThrGluAlaArgIlePheAlaIlePheValProLeuValAsnGly 139
Db      136  CTGTTACAGCAATTCACACAGACGAGTATTTCCGCCGCGTGTCCCGTCTCTGAACCTCC 195
Qy      140  LeuArgLeuValIleAsnGlyLeuSerIleSerProAsnSerMetLeuIleLysSerVal 159
Db      196  ATGAGCGCTTCTGATATGATGACTCCGCTCTACACTGATGAAGCTCTGTAAATCAAGTG 255
Qy      160  ThrArgGluGlyArgAlaGluGluLeuLeuGlySerGlyProLeuPheThrValLeuAlaLeu 179
Db      256  ACACGTGAAGGAAACCAAGAGAAATTCGAGAGAGTCCACTTATTAATGTCCTGGTGGC 315
Qy      180  LeuPheSerAlaValPhePheTrpArgIleSerProIleGlyMetIleSerLeuAlaMet 199

```

```

Db 316 CTTTCAGGCTTTAGTCTTCTGCGGTAGTCCCATCGGATGCTCTCTTCGAGT 375
Qy 200 MetCysGlyValAspGlyIleAlaAspIleMetGlyValGlyPheGlySerThrIle 219
Db 376 ATGACCGGGGCGATGCTTCCCTGACATGTTGGGAGGAGTATGGCTACCGAAGCTG 435
Qy 220 ProTyrAspProArgLysSerThrPheIleGlySerIleSerMetPheIlePheGlyPhe 239
Db 436 CCAATTCATCGAAGAGAGCTGGGCGGACATCTCATGTTCTGTTCTGTTCTG 495
Qy 240 IleSerIleAlaLeuLeuTyrTyrTyrSerSerLeuGlyTyrLeuIleMetAsnTyrGlu 259
Db 496 CTGTCGGGATGATGATGCTCTACTCTCAACCTCGGTTATGATGATGATGATG 555
Qy 260 ThrThrLeuGlnArgValAlaMetValSerMetValAlaThrValAlaGluSerLeuPro 279
Db 556 GAGGACACTGTGTACCTGCGCTGTGTGACATGACGACGACAGTGTGAGTGCCTTCT 615
Qy 280 IleThrAspGlnLeuAspAspAsnIleSerValProLeuAlaThrIleLeuAlaIleTyr 299
Db 616 GTGACCGAAGTGTGTAGATGACAAATATCTGTTCTTGGCCACCATGCTGTACCTTT 675
Qy 300 LeuSerPhe 302
Db 676 CTCTTGT 684
RESULT 14
ADXS1714
ID ADXS1714 standard; cDNA, 937 BP.
XX
AC ADXS1714;
XX
DT 21-APR-2005 (first entry)
XX
DE Plant full length insert polynucleotide seqid 26454.
XX
KV plant protectant; plant growth regulant; gene therapy; plant;
KV recombinant DNA construct; physical array; plant breeding marker;
KV cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;
KV extreme osmotic condition; pathogen tolerance; pest tolerance;
KV growth rate; cell cycle pathway; disease resistance;
KV galactomannan production; lignin production; plant growth regulator;
KV yield; plant growth; plant development; seed oil; protein yield;
KV protein content; gene; ss.
XX
OS unidentified.
XX
PN US2004034888-A1.
XX
PD 19-FEB-2004.
XX
PF 28-APR-2003; 2003US-00425114.
XX
PR 06-MAY-1999; 99US-00304517.
XX
PR 05-NOV-2001; 2001US-00985678.
XX
PA (LIUJ/) LIU J.
PA (ZHOU/) ZHOU Y.
PA (KOVA/) KOVALIC D K.
PA (SCRE/) SCREEN S E.
PA (TABAS/) TABASKA J E.
PA (CAOY/) CAO Y.
XX
PI Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JE, Cao Y;
DR WPI; 2004-180133/17.
XX
PT New recombinant DNA construct, useful for improving plant tolerance to
PT cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or
PT pests, for conferring increased resistance to plant disease, or for
PT improving yield.
XX

```

```

PS Claim 1, SEQ ID NO 26454; 15pp; English.
XX
CC The invention describes a recombinant DNA construct comprising a
CC polynucleotide consisting of a sequence encoding an amino acid sequence
CC available in electronic form from the US patent office at
CC ftp.segdata.uspto.gov/sequence.html?docid:2004034888. The polynucleotide
CC of the invention are also useful in physical arrays of molecules and as
CC plant breeding markers. The recombinant DNA construct is useful for
CC improving plant tolerance to cold, heat, drought, herbicides, extreme
CC osmotic conditions, pathogens or pests, for manipulating growth rate in
CC plant cells by modification of the cell cycle pathway, for conferring
CC increased resistance to plant disease, for producing galactomannan,
CC lignin or plant growth regulators, for increasing the rate of homologous
CC recombination in plants, for improving yield by modification of
CC photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake
CC or by providing improved plant growth and development under at least one
CC stress condition or for modifying seed oil or protein yield and/or
CC content. This sequence represents a plant full length insert
CC polynucleotide that can be used in the recombinant DNA construct of the
CC invention.
SQ Sequence 937 BP; 167 A; 261 C; 277 G; 232 T; 0 U; 0 Other;
XX
Alignment Scores:
Pred. No.: 9.17e-75 Length: 937
Score: 742.00 Matches: 152
Percent Similarity: 72.88 Conservative: 57
Best Local Similarity: 53.04 Mismatches: 62
Query Match: 48.3% Indels: 17
DB: 13 Gaps: 4
US-10-634-548-2 (1-304) x ADXS1714 (1-937)
Qy 29 SerProGlyPheLeuIle-----SerSerProCysPheIleGlyLeuThr--Gly 44
Db 34 TCCCCCACTGCGTCTGCTCTGCGCTGCGCTGCCGCCGCGCGCGCTGCGCTTCG 93
Qy 45 MetGlySerAlaThrGlnLeuArgAlaArgPheSerLeuIleSer-----59
Db 94 CTGGAGAGCTCCATGCGG-----AGGGGCTCTCTGCTGCGCGCTGCGCGACCCCGGCC 144
Qy 60 -----SerAlaValAlaThrAsnSerLeuIleHisAspValGlyAlaThrVal 75
Db 145 GTGGCGGCGCTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTG 204
Qy 76 AlaValLeuGlyGlyAlaTyrAlaLeuValLeuSerPheGluSerLeuThrLysArgAsn 95
Db 205 CTCATCACCGCGCGCGCTTACTCTTGTGCGCGCTTCGACGAGCTCACGAGCGCGG 264
Qy 96 ValIleGlnGlnSerLeuSerArgLysLeuValHisIleLeuSerGlyLeuLeuPheVal 115
Db 265 CTCATCGAAGAGAGCTGTGAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 324
Qy 116 LeuAlaTyrProIlePheSerGlySerThrGlnAlaArgTyrPheAlaAlaPheValPro 135
Db 325 TCATCTTGCCCCCTGGTGTGACAAATTCACAGAGACGAGTATTTGCGCGGTTTCCCG 384
Qy 136 LeuValaGnglyLeuArgLeuValIleAsnGlyLeuSerIleSerProAsnSerMetLeu 155
Db 385 TTCCTGAATCCTCAGAGAGCTTCTGATATATGAGACTCCGCTCTACATGAGAGAGCTCTG 444
Qy 156 IleLysSerValThrArgGlnGlyValArgAlaGluGluLeuLeuValGlyProLeuPheTyr 175
Db 445 GTAAGATGAGACACCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 503
Qy 176 ValLeuAlaLeuLeuPheSerAlaValPheThrPheArgGluSerProIleGlyMetIle 195
Db 504 GTCCGTGGCTGCTGCTTCAAGCTTTTACTCTTCTGCGGTAGTCCCATCGGATGCTC 563
Qy 196 SerLeuAlaMetMetCysGlyIleAspGlyIleAlaAspIleMetGlyValGlyPheGly 215
Db 564 TCCTTGTGATGATGAGAGAGCGGTGGAGATGTTCCGCTGACATGTTGGAGAGAGATATGGC 623

```


THIS PAGE LEFT BLANK

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - nucleic search, using frame_plus2n model

Run on: March 16, 2006, 20:05:42 ; Search time 4508 Seconds

(without alignments)
3155.118 Million cell updates/sec

Title: US-10-634-548-2

Perfect score: 1336

Sequence: 1 MAATLPLSPINHOICRFGNN.....DDNISVPLATILAAVLSRGY 304

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+g2n.model -DEV=xlh
-Q=/abs/ABSSWB.epool/US10634548/runat_16032006_084435_10357/app_query.fasta_1
-DB=EST -QFMT=fastcap -SUPFIX=g2n.rst -MINMATCH=0.1 -LOOFC=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=biom62 -TRANS=human40.cdi -LIST=45
-DOCLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTMT=ptc -NORM=ext -HAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abs02h
-USER=US10634548 @GNN_1_5315 @runat_16032006_084435_10357 -NCPU=6 -ICPU=3
-NO MMAP -NEG SCORES=0 -WAIT -DSBLOC=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=10 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:*

1: gb_est1:*\n2: gb_est2:*\n3: gb_est3:*\n4: gb_hlc:*\n5: gb_est4:*\n6: gb_est5:*\n7: gb_est6:*\n8: gb_est7:*\n9: gb_gsa1:*\n10: gb_gsa2:*\n11: gb_gsa3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	924	60.2	636	1	AU236980 AU236980
2	829	54.0	864	8	DR914339 EST110587
3	783	51.0	524	6	CB261126
4	775.5	50.5	881	7	CR286787
5	766	49.9	650	7	CO981750 GM89012A2
6	762	49.6	883	7	CV711742 UCRPT01.0
7	757	49.3	787	6	CA297624 SCCSD2C0

8	745	48.5	507	6	CB261834
9	723	47.1	734	8	DN207999
10	716	46.6	654	2	BG412580
11	701	45.6	675	8	DN229569
12	678	44.1	941	7	CR297713
13	671	43.7	623	1	AT171218
14	665	43.3	690	6	CA277256
15	658	42.8	650	4	AY109900
16	647.5	42.2	743	8	DN208011
17	622	40.5	483	1	AT955595
18	616.5	40.1	680	8	DN212311
19	602	39.2	618	6	CD405855
20	587	38.2	418	1	AU227991
21	579.5	37.7	634	3	BM080425
22	572.5	37.3	871	6	CP435509
23	567	36.9	690	1	AU056832
24	562.5	36.6	735	7	CN206873
25	551.5	35.9	660	7	CR936854
26	542.5	35.3	602	2	BR359582
27	542	35.3	569	1	AT833732
28	536.5	34.9	898	8	CX113647
29	535	34.8	732	3	BU611225
30	534.5	34.8	696	7	CR746830
31	533	34.7	830	8	CV760662
32	520	33.9	546	6	CA202669
33	517.5	33.7	508	7	CR748049
34	512	33.3	816	10	CU978431
35	510	33.2	596	6	CD392779
36	508	33.1	930	8	CX191216
37	507	33.0	817	6	CF212803
38	506.5	33.0	1125	8	DR740330
39	505.5	32.9	796	8	DN172147
40	502.5	32.7	728	6	CP429259
41	501.5	32.6	774	8	CX614964
42	492.5	32.1	817	6	CD573895
43	490.5	31.9	640	1	AW735812
44	490	31.9	462	5	BO536471
45	489.5	31.9	761	6	CP212727

ALIGNMENTS

RESULT 1
AU236980
LOCUS
DEFINITION
AU236980 RAPL15 Arabidopsis thaliana cDNA clone RAPL15-38-N13 5', mRNA sequence.

ACCESSION
AU236980.1 GI:19876149

KEYWORDS
EST.

SOURCE
Arabidopsis thaliana (chale cress)

ORGANISM
Arabidopsis thaliana

REFERENCE
Seki, M., Narusaka, M., Ishida, J., Kamiya, A., Satou, M., Nakajima, M., Akiyama, K., Enju, A., Oono, Y., Sakurai, T., Carninci, P., Kawai, J., Itoh, M., Ishii, Y., Arakawa, T., Shibata, K., Shinagawa, A., Muramatsu, M., Hayashizaki, Y. and Shinozaki, K.

Large scale analysis of Arabidopsis full-length cDNA

Unpublished (2002)

CONTACT: Motoaki Seki

Plant Functional Genomics Research Group

RIKEN Genomic Sciences Center

3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan

Tel: 81-298-36-4359

Fax: 81-298-36-9060

Email: msekic@riken.go.jp

An Arabidopsis full-length cDNA library was constructed essentially as reported previously (Seki et al., 1998). cDNA cleaved with BamHI and XhoI was ligated to modified lambdaPhiC-1 vector (Carninci et al., submitted for publication) digested with BamHI and SalI. This

clone is in a modified pBluescript vector. Please visit our web site (http://www.genetiken.go.jp/e/plan/index_e.html) for further details.

FEATURES

source

Location/Qualifiers

1..636
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/clone="RAFL15-38-N13"
/tissue_type="mixture of silique and flower"
/lab_host="DH10B"
/clone_1ib="RAFL15"
/note="Site_1: BamHI; Site_2: SalI"

ORIGIN

Alignment Scores:

Pred. No.:	6..8e-95	Length:	636
Score:	924.00	Matches:	185
Percent Similarity:	97.94	Conservative:	0
Best Local Similarity:	97.94	Mismatches:	4
Query Match:	60.24	Indels:	0
DB:	1	Gaps:	0

US-10-634-548-2 (1-304) x AUZ36980 (1-636)

```

QY      1 MetAlaAlaThrLeuProLeuSerProIleAenHlsgInleuCyArigPheglYAsnAn 20
DB      70 ATGGAGCAACCTTATCTTATCTCGAATCAATCAATGATGTCGGTTCGGAGCAAC 129
QY      21 SerLeuThrThrHlsArPheCySerProglYPhleuIleSerSerProCySphel 40
DB      130 TCTTTGAGACCTCACCGGTCGTCCTCTGCTGCTTCTGATTTCTTCTCTGTTTCAAT 189
QY      41 GlyLeuThrGlyMerGlySerAlaThrGlnleuAlaArgArSerleuIleSerSer 60
DB      190 GGTGTGACCGGAAATGGCTGCTGCTACTCACTACGTCGTCGTCGTCCTCTGATCTCTCA 249
QY      61 AlaValAlaThrAenSerleuHlsAenPValGlyAlaThrValAlaValleuGly 80
DB      250 GCGATTGGACGAAATTCGCTGTGCAATGAGTCGAGGCCCGCTGGCAATGCTTGGTGA 309
QY      81 AlaTyrAlaLeuValleuSerPhegluSerleuThrlySargAsnValIleGlnInser 100
DB      310 GCATGCGGCTGTCTTAACTTCGAGAGTCCACCAAGCGAAAGTATTCAACAGAGT 369
QY      101 LeuSerArglyLeuValHlsIleleuSerGlyLeuLeuPheValleuAlaTrpPro 120
DB      370 TTGACGCAAAACCTTGTCATATCTCACTCAAGCTGCTTTCGTTCTGCGGCGCAATC 429
QY      121 PheSerGlySerThrGlnAlaArgTyrPheAlaAlaPheValProleuValAsnGlyLeu 140
DB      430 TTCACGGAGTCCAGCGAGCTCGATATCTTGTCTTGTTCCTTACGTAAATGGCTTA 489
QY      141 ArgLeuValIleAsnGlyLeuSerIleSerProAsnSerMetleuIleIleSerValThr 160
DB      490 AGCGTTGTTATTAAAGGACTATCAATTCGCCAAATTCATCTAATATCCGTCACA 549
QY      161 ArgGlyGlyArgAlaGlnleuLeuLeuGlyProleuPheTyrValleuAlaLeuLeu 180
DB      550 AAGAGAGGAGAGAGAGAGAGAGTGTCTAAAGTCCCTTGTGTTCTACGTTACTCTTCT 609
QY      181 PheSerAlaValPhePheTrpArgIle 189
DB      610 TTCTCTGCGGANTTCTTCTGAGAGAG 636

```

RESULT 2

DR914339/c

864 bp mRNA linear EST 02-AUG-2005
LOCUS DR914339 Aquilegia cDNA library Aquilegia formosa x Aquilegia
DEFINITION pubescens cDNA clone COL102, mRNA sequence.

ACCESSION

DR914339.1 GI:71683702

KEYWORDS

EST.

SOURCE

ORGANISM

Aquilegia formosa x Aquilegia pubescens
Aquilegia formosa x Aquilegia pubescens
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; Ranunculales;
Ranunculaceae; Aquilegia.

REFERENCE

1 (bases 1 to 864)
Hodges, S.A., Rensink, W., Buell, C.R., Borevitz, J., Kramer, E.,
Nordborg, M. and Tomkins, J.

TITLE

Generation of ESTs from Aquilegia

JOURNAL

Unpublished (2005)

COMMENT

Other ESTs: EST1105879
Contact: Scott Hodges
Department of Ecology, Evolution and Marine Biology
University of California, Santa Barbara
Santa Barbara, CA 93106, USA

Tel: 805 893 7813
Fax: 805 893 4724

Email: hodges@lifesci.ucsb.edu

Sed primer: TTTTYYYYYYYYYYYYTTTTTT (where N = A, G & C).

FEATURES

source

Location/Qualifiers

1..864
/organism="Aquilegia formosa x Aquilegia pubescens"
/mol_type="mRNA"
/db_xref="taxon:338618"
/clone="COL102"
/tissue_type="mixed shoot and floral apical meristems,
flower buds, leaves and roots"
/lab_host="DH10B T1 (T1 and T5 phage resistance)"
/clone_1ib="Aquilegia cDNA library"
/note="Vector: pCMV SPORT6.1; Site_1: EcoRI, Site_2: NotI;
P2, P3, and P4 lines of Aquilegia formosa x A. pubescens
were grown from seed in greenhouses at UC Santa Barbara.
From these plants three sets of tissue were collected: 1)
Small flower buds (<10 mm) and very young inflorescences
(71 & 29% by weight respectively), 2) Medium (7-20 mm) and
large (at or near anthesis) flower buds (65 & 35% by
weight respectively) and 3) Shoot apical meristems. A
fourth set of tissue was collected from plants of A.
formosa. These plants were grown from seed in sand and at
approximately 1 month root tissue and leaf tissue of
various developmental stages were collected (84 & 16% by
weight respectively). Total RNA was extracted from each
set of tissue and pooled in the following proportions:
1.5X from sets 1 & 2, 1X from sets 3 & 4. From the pooled
total RNA, mRNA was extracted and enriched for full-length
messages and then normalized with proprietary methods by
Invitrogen."

ORIGIN

Alignment Scores:

Pred. No.:	8..31e-84	Length:	864
Score:	829.00	Matches:	156
Percent Similarity:	82.28	Conservative:	47
Best Local Similarity:	63.28	Mismatches:	44
Query Match:	54.04	Indels:	0
DB:	8	Gaps:	0

US-10-634-548-2 (1-304) x DR914339 (1-864)

```

QY      58 IleserSerAlaValAlaThrAenSerleuHlsAenPValGlyAlaThrValAlaVal 77
DB      816 ATATTAGATTCACACTGCGCAACGCGACCTTGGCAGATGACGAGTCCGCGGACATTGTT 757
QY      78 LeuGlyGlyAlaTyrAlaLeuValleuSerPhegluSerleuThrlySargAsnValIle 97
DB      756 AGTGTGTGCTTACTGCTTGTACGTTCTTGAATATCTTAACGAGAAATTTAAATC 697
QY      98 GlnGlnSerleuSerArglyLeuValHlsIleleuSerGlyLeuLeuPheValleuAla 117
DB      636 AAACAGATTTGACGAGAAAGTGTTCACATATGTCTGGTCTCTATTACATGCTTCT 637
QY      118 TrpProIlePheSerGlySerThrGlnAlaArgTyrPheAlaAlaPheValProleuVal 137

```



```

Db      636 TGGCAATTTTATGATCAACGAGGTGCTACTTTGCTGATTTTGCCTTTTG 577
Qy      138 AaaglyLeuAglLeuValIleAaaglyLeuSerIleSerProAaSerMetLeuIlelys 157
Db      576 AATGCTTGAGGCTTGCTCTCAATGACTTTCATGACGACTGAGAGGAGGCTTGCTGAAA 517
Qy      158 SerValThrArgGluGlyAaAglAglLeuLeuLeuGlyProLeuPheThrValLeu 177
Db      516 TCTGTCACACGAGGAGAAATCCAAAGAACTTCGAGAGGCTCTCTCACTACCTTCTCA 457
Qy      178 AlaLeuLeuPheSerAlaValAaPhePheTArgGluSerProIleGlyMetIleSerLeu 197
Db      456 ATGTTGATTCGACAGTCACTACTATTTCTGGGTGAGTCTCCAGTTGGATGATTCACACTA 397
Qy      198 AlaMetMetCygGlyGlyAaaglyIleAlaAaPheIleMetGlyAaGlyPheGlySerThr 217
Db      396 GCTATGATGTGTGGGGAGTGGTTTCTGATATCATATGAAAGAAATTTGGGTCAGACTA 337
Qy      218 LysIleProGlyAaAaProArgLysSerTrpAlaGlySerIleSerMetPheIlePheGly 237
Db      336 AAGATCCATATATATCAACAAAGATTTAGCTGAGAGTTCTCCATGTTTCTATTGT 277
Qy      238 PhePheIleSerIleAlaLeuLeuTyTyTyTySerSerLeuGlyTyTyLeuHisMetAa 257
Db      276 TTTCTGTTCTCAATCATGTGTGCTATATTTTTCAGCTCTTGATTTTTCATTTGAC 217
Qy      258 TrpGluThrThrLeuGluAaGlyAaAlaMetValSerMetValAlaThrValAlaGluSer 277
Db      216 TGGTCATCTACAGAGTGGAAAGTGGCTTAGAGCTTGGGCAACCATAGTGGAGTCC 157
Qy      278 LeuProIleThrAaPheGluLeuAaAaPheAaHisSerIleSerProLeuAlaThrIleLeu 297
Db      156 CTCTCTACTACAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 97
Qy      298 AlaTyLeuSerPheGlyTyTyr 304
Db      96 GCATTATGATTTTTCAGCTAT 76

```

```

RESULT 3
CB261126 524 bp mRNA linear EST 06-NOV-2003
LOCUS 26-B9406-012-002-C04-c7r MP12-ADIS-012 Arabidopsis thaliana cDNA
DEFINITION clone MP12p769C042Q 5-PRIME, mRNA sequence.
ACCESSION CB261126
VERSION CB261126.1 GI:32885899
KEYWORDS EST.
SOURCE Arabidopsis thaliana (chale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosoids; eurosoids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 524)
Schmid, K.J., Soeensen, T.R., Stracke, R., Torjek, O., Altmann, T.,
Mitchell-Olds, T. and Weisshaar, B.
Large-scale identification and analysis of genome-wide
single-nucleotide polymorphisms for mapping in Arabidopsis thaliana
Genome Res. 13 (6), 1250-1257 (2003)
JOURNAL PUBMED 1279357
COMMENT Contact: Weisshaar B
Max-Planck-Institute for Plant Breeding Research
Carl-von-Linne Weg 10, 50829 Koeln, Germany
Fax: 00492215062851
Email: weisshaar@mpiz-koeln.mpg.de
Insert Length: 524 Std Error: 0.00
Plate: 2 row: C column: 04
Seq primer: T7R: CTAATACGACTCACTATAGGA.
FEATURES
Source
1..524
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/ecotype="Landberg erecta"
/db_xref="GABI:589752"

```

```

FEATURES
Source
1..524
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/ecotype="Landberg erecta"
/db_xref="GABI:589752"

```

```

/db_xref="taxon:3702"
/clone="MP12p769C042Q"
/issue_type="whole plant"
/dev_stage="adult plant, mixed stresses"
/lab_host="E. coli Blue MRP"
/clone_id="MP12-ADIS-012"
/note="Vector: pSPORT1; Site 1: SalI; Site 2: NotI; cDNA
library from Arabidopsis thaliana, accession Landberg
erecta; six weeks old total plants grown under long-day
conditions in soil, whole adult plants were treated for 24
hours with different stresses, (1) at 4M-0 C in the dark,
(2), at 37 Grad C in the dark, (3) lying in the lab after
removing from soil, (4) in the greenhouse after wounding
leaves with a foreceps, (5) in the lab watering with a 150
mM NaCl solution, (6) at 26 Grad C in the light/UV, equal
quantities of stressed plant material were pooled; library
was made at the Max-Planck-Institute for Plant Breeding
Research, Cologne, Germany; cloning sites SalI-NotI,
primer sites and orientation:
T7-SalI-CCACGGCTCCG-5prime-cDNA-polylA-CC-NotI-SP6; Note:
Sequencing granted in the context of the GABI Arabidopsis
Verbund I: Genetic Diversity, 'Establishment of
high-efficiency SNP-based mapping tools and development of
methods for genome-wide mutation detection', PI: Bernd
Weisshaar Sequence submission managed by RZPD/GABI-Primary
database: http://gabi.rzpd.de. This clone is available
from RZPD; contact RZPD (clone.rzpd.de) for further
information."

```

ORIGIN

```

Alignment Scores:
Prod. No.: 7 136-79 Length: 524
Score: 783.00 Matches: 157
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 51.0% Indels: 0
DB: 6 Gaps: 0

```

US-10-634-548-2 (1-304) x CB261126 (1-524)

```

Qy      1 MetAlaAlaThrLeuProLeuSerProIleAaHisGluLeuCyArgPheGlyAaAa 20
Db      52 ATGGAGAGCACTTAACCTTATCCATCCGATCATATCATGTTGTGCGGAAACAC 111
Qy      21 SerLeuThrThrHisAaGlyPheCysSerProGlyPheLeuIleSerSerProCysPhe 40
Db      112 TCTTGAACGACTCACCGGCTGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 171
Qy      41 GlyLeuThrGlyMetGlySerAlaThrGluLeuArgAlaAaGlySerLeuIleSerSer 60
Db      172 GATTGACCGGAATGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 231
Qy      61 AlaValAlaThrAaSerLeuLeuHisAaPheValGlyAlaThrValAlaValLeuGlyGly 80
Db      232 GCAGTTGAGCAAGATTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 291
Qy      81 AlaTyAlaLeuValLeuSerPheGluSerLeuThrTyAaAaAaValIleGluGlnSer 100
Db      292 GCATACGCGCTTGCTTAAAGCTTCGAGAGTCTCAACCAAGCAACGTCATTCAACAGAGT 351
Qy      101 LeuSerArgLysLeuValHisIleLeuSerGlyLeuLeuPheValIleAlaTrpProIle 120
Db      352 TTGAGCAAAAGCTTGCGATATCTCTCAGAGTCTGCTTTCGATCTTGCGTGGCCATTC 411
Qy      121 PheSerGlySerThrGluAlaArgTyPheAlaAlaPheValProLeuValAaGlyLeu 140
Db      412 TTCAAGCGATGACGAGAGCTGATATCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 471
Qy      141 ArgLeuValIleAaGlyLeuSerIleSerProAaSerMetLeuIlelys 157
Db      472 AGGCTTGTTATTAACGACTATCATTTCCCAAAATTCAGATCAATCAAA 522

```

RESULT 4

CK286787 881 bp mRNA linear EST 02-AUG-2004
 LOCUS CK286787
 DEFINITION Nicotiana benthamiana mixed tissue cDNA library, normalized, full-length Nicotiana benthamiana cDNA clone NEMB462 5' end, mRNA sequence.
 ACCESSION CK286787.1 GI:39862689
 VERSION CK286787
 KEYWORDS EST.
 SOURCE Nicotiana benthamiana
 ORGANISM Nicotiana benthamiana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; asterids; lamids; Solanales; Solanaceae; Nicotiana.
 REFERENCE 1 (bases 1 to 881)
 AUTHORS Buell, C.R., Hart, A., Zisman, V., Karaymicheva, S.A., Day, B., Staskavicz, B., Jin, H., and Baker, B.
 TITLE Generation of EST sequences from Nicotiana benthamiana
 JOURNAL Unpublished (2003)
 COMMENT Other ESTs: EST749510
 Contact: Robin Buell
 The Institute for Genomic Research
 9712 Medical Center Dr, Rockville, MD 20850, USA
 Email: potato-array@tigr.org
 Clones can be requested from the University of Arizona Genomics Institute via http://genome.arizona.edu/orders/
 Seq primer: ATT TAG GTG ACA CTA TAG.
 FEATURES
 source
 location/Qualifiers
 1..881
 /organism="Nicotiana benthamiana"
 /mol_type="mRNA"
 /db_xref="taxon:4100"
 /clone="NEMB462"
 /tissue_type="abiotic and biotic stress-treated leaves, callus tissue and root tissue"
 /lab_host="DH10B-Tona"
 /clone_lib="Nicotiana benthamiana mixed tissue cDNA library, normalized, full-length"
 /note="Vector: pCMVSPORT6.1; Site 1: EcorI; Site 2: NotI; supplier: RNA was isolated from Nicotiana benthamiana tissues that include callus, roots from liquid culture grown plants, heat-stressed leaves (38 C, 3 hr and 6 hr), cold-stressed leaves (5 C 3 hr, 6hr), and pathogen challenged leaves (Pseudomonas syringae pv. tomato 12 hr; Pseudomonas syringae pv. campestris 12 hr, 18hr; Pseudomonas syringae pv. phaseolicola 18hr, and Pseudomonas campestris pv. vesicatoria 18hr). RNA was isolated from these tissues and pooled in approximately equal molar amounts."
 ORIGIN
 Alignment Scores:
 Pred. No.: 1,15e-77 Length: 881
 Score: 775.50 Matches: 150
 Percent Similarity: 73.3% Conservative: 43
 Best Local Similarity: 57.8% Mismatches: 51
 Query Match: 50.5% Indels: 23
 Gaps: 5
 US-10-634-548-2 (1-304) x CK286787 (1-881)
 Oy 16 ArgpHeGlyAsn-----AasnerLeuThrThrHis----- 25
 Db 82 CcTtTtCGAATCGCTTCCTCTCTCATCTGATACCCACACTTACATTGCGCGA 141
 Oy 26 ArgpHeGlySerPro-----GlypHeLeuIleSerSerProCybHeIleGlyLeuThr 43
 Db 142 CcATTCTCTCCGCCGACAGTTCCCGAGCGTTCCTGCTGCCGACCGCGTTTAA 201
 Oy 44 GlyMetGlySerAlaThrGlnLeuArgAlaArgArgSerLeuIleSerSerAlaValAla 63
 Db 202 GcTGTTTGAATGCT-----ATAATTTGGAATGCG----- 231
 Oy 64 ThrAsnSerLeuLeuHisAspValGlyAlaThrValAlaValLeuGlyGlyAlaTyrAla 83

Db 232 ---GGCTCAGTGTGACAGATGCGGAGCCACGGCTTTGGTCATCGCTGAGCTTACCCC 288
 Oy 84 LeuValLeuSerPheGlySerLeuThrLysArgAsnValIleGlnGlnSerLeuSerArg 103
 Db 289 CTGTTCCTACTTTCGATTTCTTATCCAGCGCAAGCTCATTTAGACAAATTTAGCGA 348
 Oy 104 LysLeuValHisIleLeuSerGlyLeuLeuPheValLeuAlaTyrProIlePheSerGly 123
 Db 349 AAGCTTGCCACATTTGTCTGCTGCTTTTATGCTTCTGCGCAATTTTACGCA 408
 Oy 124 SerThrGlnAlaArgTyrPheAlaAlaPheValProLeuValAsnGlyLeuArgLeuVal 143
 Db 409 TCAAGATGGGACCCCTACTTGTCTGTAGTTCGCTTACAACTGTTTGAAGACTTGTG 468
 Oy 144 IleAsnGlyLeuSerIleSerProAsnSerMetLeuIleLysSerValThrArgGlyGly 163
 Db 469 ATTATGCGCTTTCCTTGGCTACTGATGAGGAGCTTTAAATTTGTTACTGCGGAGGA 528
 Oy 164 ArgAlaGlyGlyLeuLeuLeuGlyProLeuPheTyrValLeuAlaLeuLeuPheSerAla 183
 Db 529 AAGCCAGAGAAATTCCTTAGAGGCGCTTATATATGTTAGTTAATTTTACGCCA 588
 Oy 184 ValPhePheTyrPArgLysSerProIleGlyMetIleSerLeuAlaMetMetCysGlyGly 203
 Db 589 CTTCCTTTTGGCGTGAATCACAGTTGAGACTAATTCGTTAGCAATGATGTGTGTGT 648
 Oy 204 AapGlyIleAlaAapIleMetGlyArgLysPheGlySerThrLysIleProTyrAsnPro 223
 Db 649 GATGGAATTCGTGATATGTTGGAGAAAGTTGGGTCTCATAAATATCCCTTATATAA 708
 Oy 709 CAGAAATTTGGTGTGATGAGCTCTCCATGTTTGTTCGGTTTCCTGCTGATGCGG 768
 Db 244 LeuLeuTyrTyrTyrSerSerLeuGlyTyrLeuHisMetAsnTyrGlnThrThrLeuGln 263
 Oy 769 ATGCTCTTATCTTCTCTCTCTGGAATCTTCACTTGAGTTGGTTTCAACTGTAA 828
 Db 264 ArgValAlaMetValSerMetValAlaThrValAlaGlySerLeuProIle 280
 Oy 829 AGTGAAGCTGTGGTGTCTATTATAGCACTATGTTGAGTCTTCACTTATTT 879
 RESULT 5
 CO981750/c 650 bp mRNA linear EST 13-SEP-2004
 LOCUS CO981750
 DEFINITION GM89012A2C07.r1 Gm-r1089 Glycine max cDNA clone Gm-r1089-4334 3, mRNA sequence.
 ACCESSION CO981750
 VERSION CO981750.1 GI:51337884
 KEYWORDS EST.
 SOURCE Glycine max (soybean)
 ORGANISM Glycine max (soybean)
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
 REFERENCE 1 (bases 1 to 650)
 AUTHORS Vockin, L., Shoemaker, R., Keim, P., Polacco, J.C., Retzel, E., Khanna, A., Shealy, R., Clough, S., Thibaut-Nissen, F., Correll, V., Erpelid, J., Gonzalez, D.O., Stromvik, M., Rodriguez-Huete, A.M., Schweitzer, P., Gong, G., and Liu, L.
 TITLE A Functional Genomics Program for Soybean (NSF 9872565) (2004)
 JOURNAL Unpublished (2004)
 COMMENT Other ESTs: BQ252454 corresponding to Gm-cl052-5684 (5')
 Contact: Vockin, L.O., PI, A Functional Genomics Program for Soybean (NSF 9872565)
 Lewin, H. A., Director, Keck Center for Comparative and Functional Genomics
 University of Illinois
 Edwin R. Madigan Building, 1201 W. Gregory, Urbana, IL 61801, USA
 Tel: (217) 244-6147
 Fax: (217) 333-4582

Email: l-vodkin@iuc.edu
 Plate: GM69012A2 row: C column: 07
 Seq primer: 5'-TTTTTTTTTTTTTTTTT(A/C/G)-3'
 High quality sequence stop: 650.
 Location/Qualifiers

FEATURES

SOURCE

1..650
 /organism="Glycine max"
 /mol_type="mRNA"
 /db_xref="taxon:3847"
 /clone_lib="Gm-r1089-4334"
 /note="The library Gm-r1089 is a sequence-driven, reracked set of 9,216 low redundancy clones selected from 38 different cDNA libraries constructed from various tissues and stages of development of soybean including 973 cDNAs from germinating cotyledons (source library Gm-c1069, Gm-c1076, and Gm-c1077); 1,465 cDNAs from various tissue and organ systems of the adult plant; 476 cDNAs from adult stem tissue (source library Gm-c1062); 1340 cDNAs from tissue culture derived somatic embryos (source libraries Gm-c1036 and Gm-c1075); 2918 cDNAs from hypocotyls or young seedlings; 742 cDNAs from germinating seedlings, shoot tips, or leaves exposed to various stresses (source libraries Gm-c1065, Gm-c1066, and Gm-c1068); 839 cDNAs from young leaves or hypocotyls exposed to bacterial and fungal pathogens (source libraries Gm-c1072, Gm-c1073, Gm-c1074, and Gm-c1084); and 463 from roots of young plants grown in hydroponic media without phosphate (source library Gm-c1087). The 5' ESTs of the source clones from the different libraries were used to select singletons, or a representative of each contig, which were reracked to form library Gm-r1089 and the cDNA clones of the reracked Gm-r1089 library were then sequenced at the 3' end. The plant genome project #9872565
 (http://soybean.genomics.croptsci.iuc.edu/) as part of creation of a low redundancy soybean cDNA set. The source cDNA libraries were constructed by the laboratories of Lila Vodkin, University of Illinois, Randy Shoemaker, Iowa State University, and Paul Keim, Northern Arizona University as part of the Public EST project, http://129.186.26.94/soybeanest.html. The contig analysis to select unique genes was performed by the laboratory of Ernest Retzel, Center for Computational Genomics and Bioinformatics, University of Minnesota, http://web.ahc.umn.edu/biodata/nstoy/. Reracking and 3' sequencing were conducted by services of the University of Illinois Keck Center for Comparative and Functional Genomics http://www.biotech.iuc.edu/keck.htm. Note: The corresponding 5' EST from each clone in the Gm-r1089 library is listed in the 'OTHER EST' field. The detailed information on the source library for each clone can also be obtained by referring to the clone ID of the original cDNA library that is also listed under 'OTHER EST'."

ORIGIN

Alignment Scores:

Pred. No.: 8.81e-77 Length: 650
 Score: 766.00 Matches: 145
 Percent Similarity: 84.9% Conservative: 35
 Best Local Similarity: 68.4% Mismatches: 32
 Query Match: 49.9% Indels: 0
 DB: 7 Gaps: 0

US-10-634-548-2 (1-304) x CO981750 (1-650)

QY 91 LeuThlyrARgAsnValIleGInGInserLeuSerArglyLeuValHisIleLeuSer 110
 DB 645 CTCACCTCGGAGAAACATCTCCAAACGGCTCTGAGACGAAACCTGCTCATATTATTACT 586
 QY 111 GlyLeuLeuPheValIleuAlaTrpProIlePheSerGlySerThrGlnAlaArgTyrPhe 130
 DB 585 GCGCTGCTTTTCTGCTTCTGCGCTATTTCAGCAACTCCCTAAGGCTCGTACTCTT 526

QY 131 AlaAlaPheValProLeuValAsnGlyLeuArgLeuValIleAsnGlyLeuSerIleSer 150
 DB 525 GCGCTTTTGTCCTCTGTCATATGCTTGAAGGCTTTGGTCAAGGCTCTCAGCTGCT 466
 QY 151 ProAsnSerMetLeuIleLeuSerValThrArgGlnGlyArgAlaGlnGlyLeuLeuLys 170
 DB 465 TGTATGAGGAGCATCATCAATCCGTCCAGAGAGAGAGATCCATGAATGTCTAAGG 406
 QY 171 GlyProLeuPheTyrValIleuAlaLeuLeuPheSerAlaValPhePheTyrArgGlySer 190
 DB 405 GGTCCCTTTTATGTTGTCATCTGATTTATCTGCTCTGCTGCTGCGGTAGTCT 346
 QY 191 ProIleGlyMetIleSerLeuAlaMetMetCysGlyGlyArgGlyIleAlaAspIleMet 210
 DB 345 CCAATGTGTGATCTCTTGTGCATGATGATGATGATGATGATGATGATGATGATGAT 286
 QY 211 GlyArgLysPheGlySerThrLysIleProTyrAsnProArgLysSerTyrAlaGlySer 230
 DB 285 GGTAGAAATATGAGTCCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 226
 QY 221 IleSerMetPheIlePheGlyPhePheIleSerIleAlaLeuLeuTyrTyrIleSerSer 250
 DB 225 ATGCTATGCTGCTGCTTGTGATCTTGTGATCTTGTGATGATGATGATGATGATGAT 166
 QY 251 LeuGlyTyrLeuHisMetAsnTyrGlnThrThrLeuGlnAlaGlyValAlaMetValSerMet 270
 DB 165 TTAGCAATGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 106
 QY 271 ValAlaThrValAlaGlyLeuPheProIleThrArgGlnLeuAspAspAsnIleSerVal 290
 DB 105 GTGGCAATATGAGTGTGATCTCTTCCCATTTACTAAGTAGAGTAGAGTAGAGTAGAGTAG 46
 QY 291 ProLeuAlaThrIleLeuAlaAlaTyrLeuSerPhe 302
 DB 45 CCACAGCTACCAAGGACAGTGGCATTTTTCATCTTC 10

RESULT 6

CV711742/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

ORGANISM

SOURCE

Poncirus trifoliata

Poncirus trifoliata

Poncirus trifoliata

Poncirus trifoliata

Poncirus trifoliata

Poncirus trifoliata

Poncirus trifoliata

Poncirus trifoliata

Poncirus trifoliata

Poncirus trifoliata

Poncirus trifoliata

Poncirus trifoliata

Poncirus trifoliata

Poncirus trifoliata

Poncirus trifoliata

Poncirus trifoliata

Poncirus trifoliata

Poncirus trifoliata

Poncirus trifoliata

Poncirus trifoliata

Poncirus trifoliata

Poncirus trifoliata

Poncirus trifoliata

Poncirus trifoliata

Poncirus trifoliata

Poncirus trifoliata

FEATURES

SOURCE

Location/Qualifiers

1..883

/organism="Poncirus trifoliata"

/mol_type="mRNA"

/cultivar="Pomeroy OP"

/db_xref="taxon:37690"

/clone_lib="PT_65a0016G05"

/tissue_type="Phloem"

/dev_stage="10 - 30 cm shoots"

Alignment Scores:

Pred. No.:	1,276-75	Length:	787
Score:	757.00	Matches:	137
Percent Similarity:	81.2%	Conservative:	49
Best Local Similarity:	59.8%	Mismatches:	43
Query Match:	49.3%	Indels:	0
	6	Gaps:	0

US-10-634-548-2 (1-304) x CA297624 (1-787)

```

Qy 74 ThrValAlaValLeuGlyGlyValAlaValLeuSerPheGluSerLeuThrlys 93
Db 776 ACGGTCATCATCAGTGGCCCATCTTTGTGGCGGCTTCGACGAGCTCACCGAG 717
Qy 94 ArgAenValIleGlnGlnSerLeuSerArgValLeuValHisIleLeuSerGlyLeu 113
Db 716 CGGACGCTCATGAAAGATTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 657
Qy 114 PheValLeuAlaTrpProIlePheSerGlySerThrGluAlaArgTyPheAlaAlaPhe 133
Db 656 TTCATGTCATCTGGCCCATCTTCAGTAATTCAGACAGACAGGATATTCGCCGACGTT 597
Qy 134 ValProLeuValIleGlnGlnSerLeuSerArgValIleGlnGlyLeuSerProIle 153
Db 596 GTTCATCTCTGAACTCATAGGCTTCGACATAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 537
Qy 154 MetLeuIleLeuSerValThrArgGluGluValArgAlaGluLeuLeuLeuGlyProLeu 173
Db 536 GCTCAGTAAATCATGATGACAGTAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 477
Qy 174 PheTyValLeuAlaLeuLeuPheSerAlaPhePheTrpArgGluSerProIleGly 193
Db 476 TATTATGCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 417
Qy 194 MetIleSerLeuAlaMetMetCysGlyGlyAspGlyIleAlaAspIleMetCysArglys 213
Db 416 ATTGATCTCTGTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 357
Qy 214 PheGlySerThrIleProTyranProArgIleSerTrpAlaGlySerIleSerMet 233
Db 356 TATGCTCAGTAAAGCTCCATTCATGATGATGATGATGATGATGATGATGATGATGATGAT 297
Qy 234 PheIlePheGlyPhePheIleSerIleAlaLeuLeuTyTyTySerSerLeuGlyTy 253
Db 296 TTCATTTCTGTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 237
Qy 254 LeuHisMetAenTrpGluThrThrLeuGlnArgValAlaMetValSerMetValAlaThr 273
Db 236 ATTCATGTTATCTGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 177
Qy 274 ValValGluSerLeuProIleThrAspGlnLeuAspAspAsnIleSerValProLeuAla 293
Db 176 GTAGGAGGATCATCTCTGTCAGTAAAGTGTGATGATGATGATGATGATGATGATGATGAT 117
Qy 294 ThrIleLeuAlaAlaTyLeuSerPhe 302
Db 116 ACCATGCTGTAAGCTTCTCTGTTT 90

```

RESULT 8
LOCUS CB261834 507 bp mRNA linear EST 06-NOV-2003
DEFINITION 85-B8864-008-010-J22-pB12-MP12-ADIS-008 Arabidopsis thaliana cDNA
clone MP12p7J2210Q 5-PRIME, mRNA sequence.

ACCESSION CB261834.1 GI:32886607
VERSION CB261834
KEYWORDS EST.
SOURCE Arabidopsis thaliana (chale crese)
ORGANISM Arabidopsis thaliana

REFERENCE 1 (bases 1 to 507)
AUTHORS Schmid,K.J., Soerensen,T.R., Stracke,R., Torjek,O., Altmann,T.,

TITLE Mitchell-Olds,T. and Weishaar,B.
JOURNAL Large-scale identification and analysis of genome-wide
PUBMED single-nucleotide polymorphisms for mapping in Arabidopsis thaliana
12799357 Genome Res. 13 (6), 1250-1257 (2003)

COMMENT Contact: Weishaar B
ADIS DNA core facility at MP12
Max-Planck-Institute for Plant Breeding Research
Carl-von-Linne Weg 10, 50829 Köln, Germany
Fax: 00492215062851
Email: weishaar@mplz-koeln.mpg.de
Insert Length: 507 Std Error: 0.00
Plate: 10 row: J column: 22
Seq primer: PB12; GGTGGCGCGCGCTCAG.
Location/Qualifiers

FEATURES

source
1..507
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/ecotype="C24"
/db_xref="GABI:601198"
/db_xref="taxon:3702"
/clone="MP12p7J2210Q"
/tissue_type="seedling"
/dev_stage="few days old seedlings"
/lab_host="E. coli XL1-Blue MRP"
/clone_id="MP12-ADIS-008"
/note="Vector: pBluescript SK (-); Site 1: EcoRI; Site 2: XhoI; cDNA library from Arabidopsis thaliana, accession C24; seedling; Lambda ZAPII phage library was made at the Max-Planck-Institute of Molecular Plant Physiology, Golm, Germany and mass-excised at the Max-Planck-Institute for Plant Breeding Research, Cologne, Germany; cloning sites EcoRI-XhoI; Note: Sequencing granted in the context of the 'Arabidopsis Verbund I: Genetic Diversity', 'Establishment of high-efficiency SNP-based mapping tools and development of methods for genome-wide mutation detection'; PI: Bernd Weishaar Sequence submission managed by RZPD/GABI-Primary database: <http://gabi.rzpd.de> This clone is available from RZPD; contact RZPD (clone@rzpd.de) for further information."

ORIGIN

Alignment Scores:

Pred. No.:	1,536-74	Length:	507
Score:	745.00	Matches:	145
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	48.5%	Indels:	0
	6	Gaps:	0

US-10-634-548-2 (1-304) x CB261834 (1-507)

```

Qy 160 ThrArgGluGlyArgAlaGluGluLeuLeuGlyProLeuPheTyValLeuAlaLeu 179
Db 2 ACAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 61
Qy 180 LeuPheSerAlaValPhePheTrpArgGluSerProIleGlyMetIleSerLeuAlaMet 199
Db 62 CTTTCTCTGCGGTTTCTTCTGAGAGAGCTCTCAATCGGTATGATCTCGTTAGCAATG 121
Qy 200 MetCysGlyGlyAspGlyIleAlaAspIleMetGlyArgLysPheGlySerThrLysIle 219
Db 122 ATGTGTGTGGCAGTAAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 181
Qy 220 ProTyranProArgLysSerTrpAlaGlySerIleSerMetPheIlePheGlyPhePhe 239
Db 182 CTTTACACCAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 241
Qy 240 IleSerIleAlaLeuLeuTyTyTySerSerLeuGlyTyLeuHisMetAenTrpGlu 259
Db 242 ATTCATCGCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCAT 301
Qy 260 ThrThrLeuGlnArgValAlaMetValSerMetValAlaThrValValGluSerLeuPro 279

```


sequence. For reasons that are not understood many of the clones in this library lack an XhoI site at their 3' ends."

ORIGIN

Alignment Scores:

Pred. No.:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
Score: 701.00	675	127	44	34	0	0
Percent Similarity: 83.4%						
Best Local Similarity: 62.0%						
Query Match: 45.6%						

US-10-634-548-2 (1-304) x DN229569 (1-675)

```

Qy 98 GlnGlnSerLeuSerArgLyLeuValHisIleLeuSerGlyLeuLeuPheValLeuAla 117
Db 673 GAAAGAGCTTGAAGAGAGGTTGTGACGCTGCTCCGCTCTTTCATGTCATCT 614
Qy 118 TrpProIlePheSerGlySerThrGluIaArgTyPheAlaPheValProLeuVal 137
Db 613 TGGCCCCCTGTTCAGCAATTCAGACAGACAGGTAATTCGCCGCCGCTTCCCTCTG 554
Qy 138 AsnGlyLeuArgLeuValIleAsnGlyLeuSerIleSerProAsnSerMetLeuIleLys 157
Db 553 AACTCCATGAGGCTTCTGATATATGACTCCGCTCTACATGATGAGGCTCTGTAATA 494
Qy 158 SerValThrArgGluGlyArgAlaGluLeuLeuLysGlyProLeuPheTyValLeu 177
Db 493 TCAGTGACACGCGAAGAAACCAAGAGAAATTCGAGAGGCTCACTTATATATGCTTG 434
Qy 178 AlaLeuLeuPheSerAlaValPhePheTrpArgIleSerProIleGlyMetIleSerLeu 197
Db 433 GTGCTGCTGTTTCAAGCTTTTACTTCTTGGCGTGAAGTCCCATCGGATGCTCTCTG 374
Qy 198 AlaMetMetCysGlyGlyAspGlyIleAlaAspIleMetGlyArgLyAspPheGlySerThr 217
Db 373 TCGAGTAGAGAGCGGCGGAGATGTTTGTGCTGACATTTGGGAGAGGATGCTCAGCG 314
Qy 218 LysIleProTyTrpAsnProArgLySerTrpAlaGlySerIleSerMetPheIlePheGly 237
Db 313 AAGCTGCATTCATCAATCGAAGAGAGCTGGCGCGGAGATCTCGATGTTCAATTTCTGGT 254
Qy 238 PhePheIleSerIleAlaLeuLeuTyTrpTyTrpSerSerLeuGlyTyTrpLeuHisMetAsn 257
Db 253 TTCCTGCTGCTCCGCAATGATGCTTACTTCTCAAGCTTGGGTTACATTATGATATC 194
Qy 258 TrpGluThrThrLeuGlnArgValAlaMetValSerMetValAlaThrValAlaGluSer 277
Db 193 TGGGAAGAGGACCTTGTAAGCTGGCGGCTTGTGACATGACGACAGATGATGAGTGC 134
Qy 278 LeuProIleThrAspGlnLeuAspAsnIleSerValProLeuAlaThrIleLeuAla 297
Db 133 GTTCTGTGACCGAAGTGTGATGATGACATATCTGTTCTTGGCCACCATGCTGGTA 74
Qy 298 AlaTyLeuSerPhe 302
Db 73 GCTTTCTCTTCTTT 59

```

RESULT 12

CK297713

LOCUS

DEFINITION

941 bp mRNA linear EST 02-AUG-2004
EST60427 Nicotiana benthamiana mixed tissue cDNA library,
normalized, full-length Nicotiana benthamiana cDNA clone NBMDA80 5'
end, mRNA sequence.

ACCESSION
CK297713.1 GI:39884376

VERSION

KEYWORDS

SOURCE

ORGANISM
Nicotiana benthamiana
Nicotiana benthamiana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
asterids; lamiales; Solanales; Solanaceae; Nicotiana.

REFERENCE

AUTHORS

1 (bases 1 to 941)
Buell, C.R., Hart, A., Zismann, V., Karamycheva, S.A., Day, B.,
Straskiewicz, B., Jin, H. and Baker, B.

TITLE

Generation of EST sequences from Nicotiana benthamiana

JOURNAL

Unpublished (2003)

Other ESTs: EST760428

COMMENT

Contact: Robin Buell
The Institute for Genomic Research
9712 Medical Center Dr, Rockville, MD 20850, USA
Email: potato-array@tigr.org
Clones can be requested from the University of Arizona Genomics
Institute via <http://genome.arizona.edu/orders/>.
Seq primer: ATT TAG GTG ACA CTA TAG.
Location/Qualifiers

FEATURES

source

```

1..941
/organism="Nicotiana benthamiana"
/mo_type="mRNA"
/db_xref="taxon:4100"
/clone="NBMDA80"
/tissue_type="abiotic and biotic stress-treated leaves,
callus tissue and root tissue"
/lab_host="DH10B-TorA"
/clone_lib="Nicotiana benthamiana mixed tissue cDNA
library, normalized, full-length"
/note="Vector: pCMVSPORT6.1; Site_1: EcoRI; Site_2: NotI;
supplier: RNA was isolated from Nicotiana benthamiana
tissues that include callus, roots from liquid culture
grown plants, heat-stressed leaves (38 C, 3 hr and 6 hr),
cold-stressed leaves (5 C 3 hr, 6hr), and pathogen
challenged leaves (Pseudomonas syringae pv tomato 12 hr;
Pseudomonas campestris pv campestris 12 hr, 18hr;
Xanthomonas syringae pv phaseolicola 18hr, and Xanthomonas
campestris pv vesicatoria 18hr). RNA was isolated from
these tissues and pooled in approximately equal molar
amounts."

```

ORIGIN

Alignment Scores:

Pred. No.:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
Score: 1,876-66	941	148	39	49	6	6
Percent Similarity: 67.8%						
Best Local Similarity: 51.9%						
Query Match: 44.1%						

US-10-634-548-2 (1-304) x CK297713 (1-941)

```

Qy 16 ArgPheGlyAsn-----AsnSerLeuThrThrHis----- 25
Db 120 CGTTCTCGAATCTGCTTCTCTCTCACTGATACCCACACTTACATTGGCGGA 179
Qy 26 ArgPheCysSerPro-----GlyPheLeuIleSerSerProCysPheIleGlyLeuThr 43
Db 180 CGATTCTTCTGCGGATACAGTCCGACCGCTTCTGATCTGCGGACCGCGTTTAC 239
Qy 44 GlyMetGlySerAlaThrGlnLeuArgAlaArgArgSerLeuIleSerSerAlaValAla 63
Db 240 GGTGGTTTGANGC-----ATGATTGAGATGCG----- 269
Qy 64 ThrAsnSerLeuLeuHisAspValGlyAlaThrValAlaValLeuGlyValAlaTyrrAla 83
Db 270 ---GGCTAGAGCTGACAGATGCGGAGCCACCGCTTGTGTCATGCTGGTCTTACGCC 326
Qy 84 LeuValLeuSerPheGlnSerLeuThrLysArgAsnValIleGlnGlnSerLeuSerArg 103
Db 327 CTGTTCACACTTTCGATTTCTTATCCAGAGCGAAGCTCATTTGAACAGATTTAAGCACA 386
Qy 104 LysLeuValHisIleLeuSerGlyLeuLeuPheValLeuAlaTrpProIlePhe----- 121
Db 387 AAGCTTTCACATATTTCTGTGCTGCTTTTATGGGCTCTGCGCAATTTTCAAGTAT 446
Qy 121 ----- 121

```


Db 447 TATTTGCTTACATATATGATTCGAAAACATATTTGAAGTTCTCACTGATGATTTAG 506
 Qy 122 -----SerGlySerThrGluAlaArgTyrPheAlaAlaPheValProLeuValAs 138
 Db 507 ATGTTAGATTGATGATCAGATGAGGACGCTACTTCTCTGATGATTCGCTTACAAA 566
 Qy 138 nGlyLeuArgLeuValIleAsnGlyLeuSerIleSerProAsnSerMetLeuIleLeuSe 158
 Db 567 CTGTTGAGACTGATGATTTATGAGCCTTTCTTTGGCTACTGATGAGGACCTTAAATTC 626
 Qy 158 rValThrArgGluGlyArgAlaGluGluLeuLeuValGlyProLeuPheTyrValLeuAl 178
 Db 627 TGTACTCGGAGAGAAAGCAGAAAGATGCTTAGAGGGCCTCATATTTATGTTCTAGT 686
 Qy 178 AleuLeuPheSerAlaValPhePheTyrArgGlySerProIleGlyMetIleSerLeuAl 198
 Db 687 GTTATATTTGACCGCATCTCTCTTTTGGCGTGAAGCACAGTGAAGTAATTTCTTATGC 746
 Qy 198 AMetMetCysGlyGlyArgGlyIleAlaAspIleMetGlyArgGlyPheGlySerThrLy 218
 Db 747 AATGATGTGTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 806
 Qy 218 sIleProTyrAsnProArgLySerTyrAlaGlySerIleSerMetPheIlePheGlyPhe 238
 Db 807 AATCCCTTATATATAACAGAAAGTTGGCTGTGAGCTCTCCATGTTGTTGTTGCTGTT 866
 Qy 238 ePheIleSerIleAlaLeuLeuTyrTyrTyrSerSerIleGlyTyrLeuIleMetAsnTr 258
 Db 867 CTGTGTGTCATGTCGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 926
 Qy 258 pGluThrThrLeu 262
 Db 927 GCTTCAACTGTA 939

RESULT 13
 LOCUS A1712188/c 623 bp mRNA linear EST 02-PEB-2000
 DEFINITION 60506SD08.x1 605 - Endosperm cDNA library from Schmidt lab Zea mays
 ACCESSION A1712188
 VERSION A1712188.1 GI:5006126
 KEYWORDS EST.
 SOURCE Zea mays
 ORGANISM Zea mays
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
 clade; Panicoidae; Andropogoneae; Zea.
 REFERENCE 1 (bases 1 to 623)
 AUTHORS Walbot, V.
 TITLE Maize ESTs from various cDNA libraries sequenced at Stanford
 JOURNAL Unpublished (1999)
 COMMENT Contact: Walbot V
 Department of Biological Sciences
 Stanford University
 855 California Ave, Palo Alto, CA 94304, USA
 Tel: 650 723 2227
 Fax: 650 725 8221
 Email: walbot@stanford.edu
 Plate: 605065 row: D column: 08.
 FEATURES
 SOURCE
 1. 623
 /organism="Zea mays"
 /mol_type="mRNA"
 /cultiivar="Ohio43"
 /db_xref="taxon:4577"
 /tissue_type="nucellar, embryo, and endosperm"
 /dev_stage="10-14 days post-pollination"
 /lab_host="DHS(alpha)"
 /clone_lib="605 - Endosperm cDNA library from Schmidt lab"
 /note="Organ: Kernel; Vector: PAD-GAL4-2; Site: 1: EcoRI;
 Site 2: XhoI; Kernel endosperm cDNA library from Schmidt
 lab"

ORIGIN
 Alignment Scores:
 Pred. No.: 6,28e-66 Length: 623
 Score: 671.00 Matches: 121
 Percent Similarity: 82.7% Conservative: 41
 Best Local Similarity: 61.7% Mismatches: 34
 Query Match: 43.7% Indels: 0
 DB: 1 Gaps: 0

US-10-634-548-2 (1-304) x A1712188 (1-623)

Qy 107 HisIleLeuSerGlyLeuLeuPheValLeuAlaTrpProIlePheSerGlySerThrGly 126
 Db 621 CACGTGCTATCCGGCGTCCTTTTCAATGTCATTCGGCCCTTTCAGCATTCACAGAA 562
 Qy 127 AlaArgTyrPheAlaAlaPheValProLeuValAsnGlyLeuArgLeuValIleAsnGly 146
 Db 561 GCCCGTATTTGGCGGGTGTCCCGTTCCTGAATCCATGAGGCTTGTATATATGA 502
 Qy 147 LeuSerIleSerProAsnSerMetLeuIleLeuSerValThrArgGluGlyArgAlaGlu 166
 Db 501 CTCGCTCTCACTGATGAAAGCTCTGTATAATCAGTGACACGTGAAGAAACAGAG 442
 Qy 167 GluLeuLeuValGlyProLeuPheTyrValLeuAlaLeuPheSerAlaValPhePhe 186
 Db 441 GAATTCCTGAGAGATCCATTTATATGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 382
 Qy 187 TrrArgGlySerProIleGlyMetIleSerLeuAlaMetMetCysGlyGlyAspGlyIle 206
 Db 381 TGGCGTGAATCCCGCATCGGATGATGATGATGATGATGATGATGATGATGATGATGAT 322
 Qy 207 AlaAspIleMetGlyArgGlyPheGlySerThrIlePheTyrAsnProArgLySer 226
 Db 321 GCTACATTCGCGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 262
 Qy 227 TrrAlaGlySerIleSerMetPheIlePheGlyPheIleSerIleAlaLeuLeuTyr 246
 Db 261 TGGCGCGGAGACATTCGATTTTCAATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 202
 Qy 247 TyrTyrSerSerLeuGlyTyrLeuIleMetAsnTrpGluThrTrpLeuGlnArgValAla 266
 Db 201 TACTTCCAGCCCGGGTATCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 142
 Qy 267 MetValSerMetValAlaThrValAlaGlySerLeuProIleThrAspGlnLeuAsp 286
 Db 141 CTGTGTGACATGACAGCAACAGTGTGAGTGCCTGCTGACCGAAGTGTGATGATGAC 82
 Qy 287 AsnIleSerValProLeuAlaThrIleLeuAlaAlaTyrLeuSerPhe 302
 Db 81 AACATATCTGTTCTTTGGCCACCATGCTGTAGCTTTTCTTGT 34

RESULT 14
 LOCUS CA277256/c 690 bp mRNA linear EST 26-SEP-2003
 DEFINITION SCACSD2014H08.g.SD2 Saccharum officinarum cDNA clone SCACSD2014H08
 5', mRNA sequence.
 ACCESSION CA277256
 VERSION CA277256.1 GI:35998743
 KEYWORDS EST.
 SOURCE Saccharum officinarum
 ORGANISM Saccharum officinarum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
 clade; Panicoidae; Andropogoneae; Saccharum; Saccharum officinarum
 complex.
 REFERENCE 1 (bases 1 to 690)
 Vettore, A.L., da Silva, F.R., Kemper, E.L. and Arruda, P.
 The libraries that made SUGEST
 JOURNAL Genet. Mol. Biol. 24 (1-4), 1-7 (2001)
 COMMENT Contact: Arruda P
 Centro de Biologia Molecular e Engenharia Genetica
 Universidade Estadual de Campinas

Caixa Postal 6010, 13083-970, Campinas SP, Brazil
Tel.: 55 19 3788 1137
Fax: 55 19 3788 1089
Email: parvuda@unicamp.br
Clone distribution: clone distribution information can be found
through the Brazilian Clone Collection Center (BCCC) at
<http://www.bccccenter.fcav.unesp.br>
Plate: 014 row: H column: 08
Seq primer: T7 Promoter Primer.
location/Qualifiers
1. .590

<http://sucest.lad.ic.unicamp.br/public>

US-10-634-548-2 (1-304) x CA277256 (1-690)

Oy	103	ArxlyseuValHisrlleuseSerGlyLeuLeuPheValleuAATPProIlePheSer	122
Db	689	AGAAAGCCTGTGCATGTGCTATCTGGCGCTCTGTTCAATCATCTTGGCACTGTTCAGT	630
Oy	123	GlySerThrGluAlaArgTyrPheAlaAlaPheValProLeuValAsnGlyLeuArgLeu	142
Db	629	ATTGCAACAAACACAGGTAATTCGGCGAGTGTGTTCATTCCTCACTCCATAGAGCTT	570
Oy	143	ValIleAsnGlyLeuSerIleSerProAsnSerMetIleLeuSerValThrArgGlu	166
Db	569	CTGACATACGGAAGCTCCGGCTCTCACTGATGAAAGCTTGGTAAATCATGATACACGTGAA	510
Oy	163	GlyArgAlaGlnGluLeuLeuLeuGlyProLeuPheTyrValIleuAlaLeuLeuPheSer	182
Db	509	GGAAACCAAGAGAAATTACTGAGAGAGGTCCACTCTATTATGTGTGGTCTCTCTTCAGT	450
Oy	183	AlaValIlePheThrPheArgIleSerProIleGlyMetIleSerLeuAlaMetMetCysGly	202
Db	449	GTTTACTCTTCTGGCGGTGATGTCCTCCCATTTGGATATGTAATCTTTCATATATAGAGCT	390
Oy	203	GlyAspGlyIleIleAlaAspIleMetGlyArgGlyPheGlySerThrValIleProTyrAsn	222
Db	389	GGCGATGGTTTTGCTGTGCATGTTGTTGGAGAAAGATAGTGCTCAGTGAACCTCCATTCAT	330
Oy	223	ProArgIlySerThrPheAlaGlySerIleSerMetPheIlePheGlyPhePheIleSerIle	242
Db	329	GAGAAAGAGAGTGGGGCGGAAACATCTCAATGTTCAATTCGTGTTTTCTGCTGTCTGCG	270
Oy	243	AlaLeuLeuYrTyrTyrSerSerLeuGlyTyrLeuIleMetAsnTrpGluThrThrLeu	262
Db	269	ATGATGATGTTCTACTCTTCAGAACCTCGGGTTACATTCACATGTAATCTGGGAGAGGCACTC	210
Oy	263	GlnArgValAlaMetValSerMetValAlaIleThrValGluSerLeuProIleThrAsp	282

Db 209 GGTAACTGAGCCCTGTTGGACCTGGCAGCAACAGTAGTGATGCATTCCTGAGCTGAA 150
 Qy 283 GlnLeuAspAspAsnIleSerValProLeuAlaThrIleLeuAlaAlaIleValLeuSerPhe 302
 Db 149 GTTGTAGATGACAATATATCTGTTCTTGGCCACCAATGCTGGAGCTTTTCTCTTGT 90

RESULT	15
AY109900/c	
LOCUS	AY109900
DEFINITION	Zea mays CT5041_1 mRNA sequence.
ACCESSION	AY109900
VERSION	AY109900.1
KEYWORDS	GI :21213810
SOURCE	HTC.
ORGANISM	Zea mays
	Zea mays
	650 bp mRNA linear HTC 25-FEB-2005

JOURNAL Submitted (25-APR-2002) Maize Mapping Project, University of Missouri, Columbia, MO 65211, USA

COMMENT If you are interested in getting corresponding physical clones, please contact me. I have a good number of clones available.

FEATURES	Location/Qualifiers
source	1. .650

ORIGIN	
Alignment Scores:	
Pred. No.:	2,07e-64
Score:	658.00
Percent Similarity:	81.6%
Best Local Similarity:	60.7%
Query Match:	42.8%
DB:	4
Length:	650
Matches:	119
Conservative:	41
Mismatches:	36
Indels:	0
Gaps:	0

US-10-634-548-2 (1-304) X AY109900 (1-650)

```
Oy 107 HisIleLeuSerGlyLeuLeuPheValIleuAlaTrpProIlePheSerGlySerThrGlu 126
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 650 CACGCTGATTCGGGGCTCTCTGTCATGTCATCTTGCCCCCTTCACGCAATTGACAGAA 591
Oy 127 AlaArgTyrPheAlaIlePheValProLeuValAsnGlyLeuArgLeuValIleAsnGly 146
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 590 GCCCGGTATTTCCGGCGGTGTCTCCGTTCTGAACTCCATGAGGCTTCGTATATATGGA 531
Oy 147 LeuSerIleSerProAsnSerMetLeuIleIleYserValThrArgGluIleValArgAlaGlu 166
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 530 CTCCGCTCTACACAGTAAAGCTCTGTAATAATCAGTGACAGTGAAAGMAAACACAGAG 471
Oy 167 GluLeuLeuValGlyProLeuPheTyrValIleuAlaLeuLeuPheSerAlaValPhePhe 186
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 470 GAATTCGTGAGAGGTCCACTCATTAATGCTCTGCTGCTGCTGCTTCAAGCTTTAGCTTC 411
Oy 187 TrpArgGluSerProIleGlyMetIleSerLeuAlaMetMetCysGlyValAspGlyIle 206
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 410 TGGCGTAGTNNNNNATCGGATCGCTCTCTGCGATGATGAGCGGTGGCATGATGTTT 351
Oy 207 AlaAspIleMetGlyArgLeuPheGlySerThrIleValProTyrAsnProArgIleSer 226
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 350 GCTGACATTTGCGGAGGAGGATGCTCAGCGAAGCTGCCATTCAATCGAAGAAAGAGC 291
Oy 227 TrpAlaGlySerIleSerMetPheIlePheGlyPhePheIleSerIleAlaLeuLeuTyr 246
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 290 TGGCCCGGAGCATCTCATGTTCTTCTGCTTCTGCTGCTGCTGCGCATGATGATGCTC 231
Oy 247 TyrTyrSerSerLeuGlyTyrIleuHisMetAsnTrpGluThrThrLeuGlnArgValAla 266
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 230 TACTTCTCAAGCTCGGTACATGATGTATCTGGAAAGAGCACTTGTTAGCTGGCG 171
Oy 267 MetValSerMetValAlaThrValValGluSerLeuProIleThrAspGlnLeuAspAsp 286
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 170 CTTGTTGCACTAGCAGCAACAGTACGTGAGTCCCTGTGACCGAAGTTGTAGATGAC 111
Oy 287 AsnIleSerValProLeuAlaThrIleLeuAlaIleAlaTyrLeuSerPhe 302
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 110 AACATATCTGTTCTTTGGCCACAGCTGAGTACTTTCCTTGTTT 63
```

Search completed: March 16, 2006, 23:29:46
Job time : 4517 secs

THIS PAGE LEFT BLANK

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: March 16, 2006, 22:17:48 ; Search time 217 Seconds
(without alignments)
2490.225 Million cell updates/sec

Title: US-10-634-548-2

Perfect score: 1336

Sequence: 1 MAATLPLSPIHQLCRFGNN.....DDNISVPLATILAVLSRGY 304

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+.p2n.model -DEV=xlh
-Q=/abs/ABSSWB_epool/US10634548/runat_16032006_084438_10410/app_query.fasta_1
-DB=Isauesd_Patents_NA -QFMT=faeap -SUFFIX=p2n.rn1 -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR=score -HEAPSIZE=500 -MINLEN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pcio -NORM=ext -THREADS=100 -THR_MIN=0
-HOST=ad8800h -USER=US10634548 -CGEN_1_1_193 -runat_16032006_084438_10410
-NCPU=6 -ICPU=3 -NO_WMAP -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=130 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents_NA.*

1: /cgn2_6/ptodata/1/ina/1_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/H_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
7: /cgn2_6/ptodata/1/ina/PP_COMB.seq:*
8: /cgn2_6/ptodata/1/ina/RE_COMB.seq:*
9: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	118.5	7.7	1664976	3	US-08-916-421B-1 Sequence 1, Appl1
2	118.5	7.7	1664976	3	US-09-692-570-1 Sequence 1, Appl1
3	105.5	6.9	1495	3	US-09-248-796A-2050 Sequence 2010, Ap
4	105.5	6.9	1726	3	US-09-605-703B-491 Sequence 489, App
5	105.5	6.9	2187	4	US-09-605-703B-489 Sequence 3131, Ap
6	100.5	6.5	858	3	US-09-328-352-3131 Sequence 37, Appl
7	100.5	6.5	891	3	US-09-540-236-198 Sequence 5219, Ap
8	100.5	6.5	99629	3	US-09-596-002-37 Sequence 5219, Ap
9	99	6.4	870	3	US-09-489-039A-5219 Sequence 5219, Ap

10	99	6.4	1548	3	US-09-328-352-964
11	98.5	6.4	1152	3	US-09-710-279-2167
12	98.5	6.4	1359	3	US-09-134-001C-1098
13	98.5	6.4	3108	3	US-09-710-279-3427
14	97.5	6.3	903	3	US-09-543-681A-282
15	97	6.3	1227	3	US-09-543-681A-282
16	97	6.3	1296	3	US-09-549-848B-5
17	97	6.3	1266	3	US-09-688-069-5
18	97	6.3	1341	3	US-09-134-000C-3049
19	97	6.3	1383	3	US-09-902-540-8008
20	97	6.3	7021	3	US-09-902-540-8008
21	96	6.2	993	3	US-09-489-039A-3797
22	95.5	6.2	1830121	3	US-09-557-884-1
23	95.5	6.2	1830121	3	US-09-643-990A-1
24	95.5	6.2	1830121	3	US-10-158-865-1
25	95	6.2	288	3	US-09-315-294A-3968
26	95	6.2	1442	2	US-08-468-763-18
27	95	6.2	1442	2	US-08-393-996A-18
28	95	6.2	2051	2	US-08-672-814D-1
29	95	6.2	2051	2	US-09-333-696-1
30	95	6.2	2051	3	US-09-282-218A-1
31	95	6.2	2103	3	US-09-282-218A-13
32	94.5	6.2	1275	3	US-09-134-001C-2172
33	94.5	6.2	1335	3	US-09-540-236-73
34	94.5	6.2	3840	3	US-09-710-279-4013
35	94.5	6.2	4235	3	US-09-710-279-4013
36	94.5	6.2	89047	3	US-09-596-002-34
37	94	6.1	756	3	US-09-134-000C-919
38	94	6.1	1639	3	US-09-328-352-568
39	94	6.1	3230	3	US-08-961-527-203
40	93.5	6.1	630	3	US-09-583-110-414
41	93.5	6.1	682	3	US-09-107-433-1052
42	93.5	6.1	1230	3	US-09-710-279-2001
43	93.5	6.1	1230	3	US-09-710-279-2305
44	93.5	6.1	1245	3	US-09-489-039A-3698
45	93	6.1	927	3	US-09-328-352-2966

ALIGNMENTS

RESULT 1
US-08-916-421B-1
Sequence 1, Application US/08916421B
Patent No. 6503729
GENERAL INFORMATION:
APPLICANT: Bull et al.
TITLE OF INVENTION: Complete Genome Sequence of the Methanogenic Archaeon, Methanococcus jannaschii
FILE REFERENCE: PB275
CURRENT APPLICATION NUMBER: US/08/916,421B
CURRENT FILING DATE: 1997-08-22
PRIOR APPLICATION NUMBER: US 60/024,428
PRIOR FILING DATE: 1996-08-22
NUMBER OF SEQ ID NOS: 3
SOFTWARE: PatentIn version 3.1
SEQ ID NO 1
LENGTH: 1664976
TYPE: DNA
ORGANISM: Methanococcus jannaschii
FEATURES:
NAME/KEY: misc feature
LOCATION: (28222)..(28222)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (28257)..(28258)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (84773)..(84773)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (84808)..(84808)
OTHER INFORMATION: n equals a, t, c, or g

Sequence 964, App
Sequence 2157, Ap
Sequence 1098, Ap
Sequence 3427, Ap
Sequence 282, App
Sequence 650, App
Sequence 5, Appl1
Sequence 3049, Ap
Sequence 8008, Ap
Sequence 814, App
Sequence 3797, Ap
Sequence 1, Appl1
Sequence 1, Appl1
Sequence 3968, Ap
Sequence 18, Appl1
Sequence 1, Appl1
Sequence 1, Appl1
Sequence 13, Appl1
Sequence 73, Appl1
Sequence 4013, Ap
Sequence 4041, Ap
Sequence 34, Appl1
Sequence 919, App
Sequence 568, App
Sequence 203, App
Sequence 414, App
Sequence 1052, Ap
Sequence 2001, Ap
Sequence 2305, Ap
Sequence 3698, Ap
Sequence 2966, Ap

NAME/KEY: misc_feature
LOCATION: (84812)..(84812)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (98120)..(98120)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (98159)..(98159)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (98239)..(98239)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (98266)..(98266)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (98343)..(98343)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (103998)..(103998)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (148948)..(148948)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (163385)..(163385)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (191989)..(191989)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (191995)..(191995)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (231980)..(231980)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (234187)..(234187)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (234220)..(234220)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (234814)..(234814)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (309398)..(309398)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (309418)..(309418)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (312837)..(312837)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (312993)..(312993)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (319226)..(319226)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (559167)..(559167)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (559241)..(559241)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (600992)..(600992)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (622708)..(622708)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature

LOCATION: (657081)..(657081)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (657203)..(657203)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (674435)..(674435)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (682442)..(682442)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (713652)..(713652)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (741684)..(741684)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (779455)..(779455)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (779676)..(779676)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (855539)..(855539)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (871619)..(871619)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1084830)..(1084830)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1096846)..(1096846)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1119881)..(1119881)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1130881)..(1130881)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1310988)..(1310988)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1313224)..(1313224)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1349473)..(1349473)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1349491)..(1349491)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1470091)..(1470091)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1569020)..(1569020)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1602912)..(1602912)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1603734)..(1603734)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1637998)..(1637998)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1664854)..(1664854)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature

US-08-916-421B-1

Alignment Scores:

Pred. No.:	15.1	Length:	1664976
Score:	118.50	Matches:	51
Percent Similarity:	42.0%	Conservative:	43
Best Local Similarity:	22.8%	Mismatches:	59
Query Match:	7.7%	Indels:	71
DB:	3	Gaps:	13

US-10-634-548-2 (1-304) x US-08-916-421B-1 (1-1664976)

```

Qy 92 ThrlyrAaGaaValIleGlnInSerleuSerArglySleuValHileuSergly 111
Db 393736 ACTAAAGAGGTTGAACCTTGAAGACATTTATAGCAACAACATTTAGTTTGG 393795
Qy 112 LeuLeuPheValleuAlaTrpProIlePheSerglySerThr- 125
Db 393796 GTTTTAAATAGCATTTTCAGTTTAAATTTTAAATAAACAATTAATTCATTAATGTT 393855
Qy 126 -----GluAlaArgTrpPheAlaIle-----PheValProleuVal 137
Db 393856 AGTATAGTTATTTGATCTGCCTATATATTTTATGTAAAGATATTACATACCATAGTA 393915
Qy 138 AaAnglyLeuArgleuValIleAaAnglyLeuSerIleSerProAaSerMetLeuIleLys 157
Db 393916 TCAGATTTA-----TTAAAT-----CTCTGTAAA 393939
Qy 158 SerValThrArgGluGlyArgAlaGluLeuLeuLysGlyProleuPheTrValleu 177
Db 393940 AGAGAAAAAGAGATGGA-----AAAGGAGCATATATCTTGTCTATT 393981
Qy 178 AlaLeuLeuPheSer-----AlaValPhePheTrpArgGlu 189
Db 393982 GGTATGTTAACTCTCTTAATTTTAAATGATGATATAAAGCTGATATTTT----- 394032
Qy 190 SerProIleGlyMetIleSerleuAlaMetMetCysGlyGlyAspGlyIleAlaAspIle 209
Db 394033 -----GGCATCTTGATTTGCTCTGTT-----GGGAGTTCTTTAGCTACTATA 394074
Qy 210 MetGlyArgLysPheGlySerThrLysIleProLysProArgLysSerTrpAlaGly 229
Db 394075 ATAGGCATTAGC-----GGAATAATTAATAAATAATAC-----TTTGAAAA 394116
Qy 230 SerIleSerMetPheIlePheGlyPhePheIleSerIleAlaLeuLeuTyrr----- 247
Db 394117 ACGGTTCAGGATTTTAA-----GCATTTTATCTCTGCCTCATTAATTTATTCATTT 394173
Qy 248 TyrrSerSerLeuGlyTyrrLeuHleMetAaTrpGluThrThrLeuGlnArgValAlaMet 267
Db 394174 TATGGAACCTTATGGG-----ATTTTC 394194
Qy 268 ValSerMetValAlaThrValValGlnSerleuProIleThrAspGlnLeuAspAsn 287
Db 394195 GTAGCTTAAATCTCGACATTTATGAAATTTGAAGTAAGAAAAATAGATGACAAAT 394254
Qy 288 IleSerValPro 291
Db 394255 CTCTATCTTCC 394266

RESULT 2
US-09-692-570-1
; Sequence 1, Application US/09692570
; Patent No. 6797466
; GENERAL INFORMATION:
; APPLICANT: Built et al.
; TITLE OF INVENTION: Complete Genome Sequence of the Methanogenic Archaeon, Methanococ
; Patent No. 6797466
; TITLE OF INVENTION: jannaschii
; FILE REFERENCE: PB275C1
; CURRENT APPLICATION NUMBER: US/09/692,570
; PRIOR APPLICATION NUMBER: 2003-01-14
; PRIOR FILING DATE: 2003-01-14
; PRIOR FILING DATE: 1996-08-22
; PRIOR APPLICATION NUMBER: US 08/916,421

```

```

PRIOR FILING DATE: 1997-08-22
NUMBER OF SEQ ID NOS: 20
SOFTWARE: PatentIn version 3.1
SEQ ID NO 1
LENGTH: 1664976
TYPE: DNA
ORGANISM: Methanococcus jannaschii
FEATURE:
NAME/KEY: misc feature
LOCATION: (28222)..(28222)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (28257)..(28258)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (84773)..(84773)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (84808)..(84808)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (84812)..(84812)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (98120)..(98120)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (98159)..(98159)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (98239)..(98239)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (98266)..(98266)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (98343)..(98343)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (103998)..(103998)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (163385)..(163385)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (146948)..(146948)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (191989)..(191989)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (191995)..(191995)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (231980)..(231980)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature

```

LOCATION: (234187) .. (234187)	
OTHER INFORMATION: n equals a, t, c, or g	
FEATURE:	
NAME/KEY: misc_feature	
LOCATION: (234220) .. (234220)	
OTHER INFORMATION: n equals a, t, c, or g	
FEATURE:	
NAME/KEY: misc_feature	
LOCATION: (234814) .. (234814)	
OTHER INFORMATION: n equals a, t, c, or g	
FEATURE:	
NAME/KEY: misc_feature	
LOCATION: (309398) .. (309398)	
OTHER INFORMATION: n equals a, t, c, or g	
FEATURE:	
NAME/KEY: misc_feature	
LOCATION: (309418) .. (309418)	
OTHER INFORMATION: n equals a, t, c, or g	
FEATURE:	
NAME/KEY: misc_feature	
LOCATION: (312837) .. (312837)	
OTHER INFORMATION: n equals a, t, c, or g	
FEATURE:	
NAME/KEY: misc_feature	
LOCATION: (319226) .. (319226)	
OTHER INFORMATION: n equals a, t, c, or g	
FEATURE:	
NAME/KEY: misc_feature	
LOCATION: (559167) .. (559167)	
OTHER INFORMATION: n equals a, t, c, or g	
FEATURE:	
NAME/KEY: misc_feature	
LOCATION: (559241) .. (559241)	
OTHER INFORMATION: n equals a, t, c, or g	
FEATURE:	
NAME/KEY: misc_feature	
LOCATION: (600992) .. (600992)	
OTHER INFORMATION: n equals a, t, c, or g	
FEATURE:	
NAME/KEY: misc_feature	
LOCATION: (622708) .. (622708)	
OTHER INFORMATION: n equals a, t, c, or g	
FEATURE:	
NAME/KEY: misc_feature	
LOCATION: (657081) .. (657081)	
OTHER INFORMATION: n equals a, t, c, or g	
FEATURE:	
NAME/KEY: misc_feature	
LOCATION: (657203) .. (657203)	
OTHER INFORMATION: n equals a, t, c, or g	
FEATURE:	
NAME/KEY: misc_feature	
LOCATION: (674435) .. (674435)	
OTHER INFORMATION: n equals a, t, c, or g	
FEATURE:	
NAME/KEY: misc_feature	
LOCATION: (682442) .. (682442)	
OTHER INFORMATION: n equals a, t, c, or g	
FEATURE:	
NAME/KEY: misc_feature	
LOCATION: (713652) .. (713652)	
OTHER INFORMATION: n equals a, t, c, or g	
FEATURE:	
NAME/KEY: misc_feature	
LOCATION: (741684) .. (741684)	
OTHER INFORMATION: n equals a, t, c, or g	
FEATURE:	
NAME/KEY: misc_feature	
LOCATION: (779455) .. (779455)	

```

OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (779676)..(779676)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (855539)..(855539)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (871619)..(871619)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1084830)..(1084830)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1096846)..(1096846)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1119881)..(1119881)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1130881)..(1130881)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1310988)..(1310988)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1313224)..(1313224)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1349473)..(1349473)
OTHER INFORMATION: n equals a, t, c, or g

Alignment Scores:
Pred. No.      15.1      Length:      1664976
Score:         118.50     Matches:      51
Percent Similarity: 42.0%   Conservative: 43
Best Local Similarity: 22.8% Mismatches: 59
Query Match:    7.7%     Indels:      71
DB:             3        Gaps:        13
US-10-634-548-2 (1-304) x US-09-692-570-1 (1-1664976)

QY      92  ThrlyArGaenValllleGlnSerleuSerArGlyslenuValHlbleuSerGly 111
      |||||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      393736 ACTAAAAAGGTGGAACCTTGAGAGAGATTATATGACAACAACATCATTTAGTTTGA 393795

QY      112 leuleuPheValleuAlaIleuProillePheSerGlySerThr----- 125
      ::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      393796 GTTTTAATGACATTTTCGCTTTTATATTTAAAAAACAATTATATCCATTATGTT 393855

QY      126 -----GluAlaArgTyPheAlaIa-----PheValProleuVal 137
      ::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      393856 AGTATAGTATTATGGATCGCTGCTATATTTTATATGTAATAAGATATTACATCAATAGTA 393915

QY      138 AsnGlyleuAlaGlyleuValIleAsnGlyleuSerIleSerProAsnSerMetleuIleYs 157
      ::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      393916 TCAGATTTA-----TTAAAT-----CTCTGTAA 393939

QY      158 SerValThrArgGlnGlyArgAlaGlnIleuLeuGlyGlyProleuPheTyValIleu 177
      ::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      393940 AGAGAAAAAGAGCATGCA-----AAAGAGCGCATATATCTTGTATT 393981

QY      178 AlaIleuPheSer-----AlaValPhePheTrpArgGlu 189

```



```

DB 393982 GSTATGTTAATCTCATTTAATTTAATGATATGATATAAGCTGATTTTTT----- 394032
QY 190 SerProilegylMetileSerleuAlaMetMetCysglYlylApglylAlaIaaplle 209
DB 394033 -----GGCATCTGTGGTATTTTCTGTT-----GGGATCTTCTTACTTACTATA 394074
QY 210 MetGlyArglyPheBglSerThrlyslleProTyranProArglySerTrpIaagly 229
DB 394075 ATAGGCAATTAGA---GAAAAATTAATAAATAATATC-----TTTGAAAA 394116
QY 230 SerIleSerMetPheIlePheBglPhePheIleSerIleAlaLeuLeuTyrr----- 247
DB 394117 ACGGTGAGGATTTTAA---GCATTTTATCTCTGCTCATTTAATTTATATCATTT 394173
QY 248 TyrrSerSerleuGlyTyrrleuHlIweChasnTrpGluThrThrlleuGlnArgValaIaMet 267
DB 394174 TATGGAATCTTATGGG-----ATTTTC 394194
QY 268 ValSerMetValAlaThrValValGluSerleuProIleThrApgGlnleuAapIaapIa 287
DB 394195 GTAGCTTAATCTCAGCATTTATTAATTTGTAAGTAAGAAAAATAAGAAATGATGACAT 394254
QY 288 IleSerValPro 291
DB 394255 CTCTATCTTCTCT 394266

RESULT 3
US-09-248-796A-2050
; Sequence 2050, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstein et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 2050
; LENGTH: 495
; TYPE: DNA
; ORGANISM: Candida albicans
US-09-248-796A-2050

Alignment Scores:
Pred. No.: 0.00108 Length: 495
Score: 105.50 Matches: 38
Percent Similarity: 44.5% Conservative: 35
Best Local Similarity: 23.2% Mismatches: 66
Query Match: 6.9% Indels: 25
DB: 3 Gaps: 8

US-10-634-548-2 (1-304) x US-09-248-796A-2050 (1-495)
QY 152 AenSerMetLeuIleLeuSerValThrArgGluGlyArgAlaGluGluLeuLeuLeuGly 171
DB 4 AATGTACGCTCATCTCATTTGTATTTAGACACAAAGATACATTC---TATACGGG 60
QY 172 ProLeuPheTyrrValLeuAlaLeuLeuPheSerAlaValPhePheTrpArgGluSerPro 191
DB 61 ACACCTTTTATTTGGTGGTGTGTTT---GTGCTTATGCTTATTCATAAAGAC 114
QY 192 IleGlyMetIleSerLeuAlaMetMetCysglYlylApglylAlaIaapIleMetGly 211
DB 115 ATTCTGTGTTGAGCATTTTATTTAGTTGGGACACATGCACTTCCACTTTTGGT 174
QY 212 ArglyPheGlySerThrlyslleProTyranProArg-----LysSerTrp 227

```

```

DB 175 AGACATTCGGTAA-----TATCTCACAGATTCGACGGTAAATCCITA 222
QY 228 AlaGlySerIleSerMetPheIle-----PheGlyPhe 238
DB 223 GCAAGCTCTCTGTCAGTGTGTAAACCGAGTTTGACTGCTTATTTGTGTATGGCTAT 282
QY 229 PheIleSerIleAlaLeuLeuTyrrTyrrSerSerleuGlyTyrrleuHlIweChasnTrp 258
DB 223 TTC-----ATTCCTCATATATACATTTGGAACCAACAGGAAATCTACTGAAATGCC 336
QY 259 GluThrThrLeuGlnArgValaIaMetValSerMetValAlaThrValValGluSerleu 278
DB 337 TCATCAACAAGTTAGTTACCATTTATCTTGTGATGGGATTAATCTAGTATTT 396
QY 279 ---ProIleThrAap-----GlnleuAapIaapIleSerValProLeuAlaThrIle 295
DB 397 TCTGAAGTAATGATTAGTGTGATTTAGTATGATTAACCTTAACATCCAGTTTGGATGGA 456
QY 296 LeuAlaIaTyrr 299
DB 457 ACCTGTATTTAT 468

RESULT 4
US-09-605-703B-491
; Sequence 491, Application US/09605703B
; Patent No. 6962989
; GENERAL INFORMATION:
; APPLICANT: Pompejus, Markus
; APPLICANT: Kroger, Burkhard
; APPLICANT: Schroder, Hartwig
; APPLICANT: Zelder, Oskar
; APPLICANT: Haberhauser, Gregor
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING NOVEL
; FILE REFERENCE: BGI-129CP
; CURRENT APPLICATION NUMBER: US/09/605,703B
; CURRENT FILING DATE: 2000-06-27
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: 60/142,764
; PRIOR FILING DATE: 1999-09-03
; PRIOR APPLICATION NUMBER: 60/152,318
; NUMBER OF SEQ ID NOS: 2934
; SEQ ID NO 491
; LENGTH: 1726
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
; FEATURES:
; NAME/KEY: CDS
; LOCATION: (101)..(1726)
; OTHER INFORMATION: PRXA02898
US-09-605-703B-491

Alignment Scores:
Pred. No.: 0.0084 Length: 1726
Score: 105.50 Matches: 68
Percent Similarity: 39.6% Conservative: 40
Best Local Similarity: 24.9% Mismatches: 93
Query Match: 6.9% Indels: 72
DB: 4 Gaps: 13

US-10-634-548-2 (1-304) x US-09-605-703B-491 (1-1726)
QY 6 ProLeuSerPro-----IleAenHlGlnLeuCyArgPheGlyAaapIa 20
DB 499 CCTGGACCTCTGGGTGGGCGGCGCAGACAGATCACTGCTCTGTGGTGGAGAAC 558
QY 21 SerLeuThrThrHlIaArgPheCysSer-----ProGlyPheLeu----- 33
DB 559 CATGCTTGCTCTTGGCTCTGCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 618
QY 34 ---IleSerSerPro-CysPheIleGlyLeuThrGlyMetGlySerAlaThrGlnLeuArg 52
DB 619 CATTTACGCTGCTCATTAACCTGGGATCGCG---ATGGGATCCAGCGTT----- 667

```

```

QY 52 gAlaArGsrLeuIleSerSerAlaValAlaThrAnSerLeuLHIsAPValGl 72
DB 668 -----GGCTTGACTCTCTTCGACATGGTTACCGGCTCGATATG----- 706
QY 72 yAlaThrValAlaValLeuGlYAlaATrAla----- 83
DB 707 -GGCACCCTGTCACATGTGACGCGGCTCGCTCGTGCAGATGAGATTAA 765
QY 84 -----LeuValIleuSerPheGluSerLeuThrLySArgAnValIleGlnGlnSe 100
DB 766 GCCTGTGAGGCAATGCGCAATGAAACCAACCACTATGCTTCATGATGAGGCGAC 825
QY 100 rLeuSerArgLySLeuValIleIleuSerGlyLeuPheValIleuAlaTrpPro1 120
DB 826 GATCACCCTGGGATCGTTCTAGCACTC---GGCATCATCTTTGGGATTCGCGGCGCAT 882
QY 120 ePheSerGlySerThrGlnAlaArg-----TyrPheAlaAlaPheVa 134
DB 883 GATGACTGATCTTCACCGCTACTGTTCCATCTGTGGGCAATTGGGCACTGTTGT 942
QY 134 1-----ProLeuValAsnGlyLeuAr 141
DB 943 CATCGTGGGTACGTTCTCTTCTACACAGCGCTTCATGCTGTGGGCGATGGCTGG 1002
QY 141 gLeuValIleAsnGlyLeuSerIleSerProAnSerMetLeuIleLySerValThrAr 161
DB 1003 CAAGTTATCGGTGACACCGTTGGGACGTTGAAAGCCTTGCTGGGACCACTCCAAAGC 1062
QY 161 gGluGlyArgAlaGluGluLeuLeuLySgLyProLeuPheTyrValIleuAlaLeuLeph 181
DB 1063 AAACCTCGCGCTGACCGCAACTACAGCGTTCCACACTGACGTGGGCACTGCTTAGTG-- 1120
QY 181 eSerAlaValPhePheTrpArgGluSerProIleGlyMetIleSerLeuAlaMetMeCy 201
DB 1121 -----ACTGCAATTTGGCATGCTTTCTGCAACCATAGAG-- 1153
QY 201 eGlyGlyAspGlyIleAlaAspIleMetGlyArgLySArgLySerThrLyIleProTy 221
DB 1154 -----GACGCAAGTCCGACATGATGCGGAGACGATACCGGACGATTAATCTCGCA 1206
QY 221 rAnProArgLySserTrpAlaGlySerIleSerMet 233
DB 1207 GGGACCAACCAAC-----GGTTCATCAACCATG 1234

RESULT 5
US-09-605-703B-489
/ Sequence 489, Application US/09605703B
/ Patent No. 6962989
/ GENERAL INFORMATION:
/ APPLICANT: Pompeius, Markus
/ APPLICANT: Kroger, Burkhard
/ APPLICANT: Schroder, Hartwig
/ APPLICANT: Zelder, Oskar
/ APPLICANT: Haberer, Gregor
/ TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING NOVEL
/ FILE REFERENCE: BGI-129CP
/ CURRENT APPLICATION NUMBER: US/09/605,703B
/ PRIOR FILING DATE: 2000-06-27
/ PRIOR APPLICATION NUMBER: 60/142,764
/ PRIOR FILING DATE: 1999-07-08
/ PRIOR APPLICATION NUMBER: 60/152,318
/ PRIOR FILING DATE: 1999-09-03
/ NUMBER OF SEQ ID NOS: 2934
/ SEQ ID NO 489
/ LENGTH: 2187
/ TYPE: DNA
/ ORGANISM: Corynebacterium glutamicum
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (101)..(2164)
/ OTHER INFORMATION: RXN01071

```

```

US-09-605-703B-489
Alignment Scores:
Pred. No.: 0.0124 Length: 2187
Score: 105.50 Matches: 68
Percent Similarity: 39.68 Conservative: 40
Best Local Similarity: 24.94 Mismatches: 93
Query Match: 6.9% Indels: 72
DB: 4 Gaps: 13

US-10-634-548-2 (1-304) x US-09-605-703B-489 (1-2187)
QY 6 ProLeuSerPro-----IleAnHISglnLeuCyArgPheGlyAsnAn 20
DB 499 CTGGAGACCTTGAGTGCGCGCGGACAGACATGACCTGCTGTGTGTGAGGAAAC 558
QY 21 SerLeuThrThrIleArgPheCySer-----ProGlyPheLeu----- 33
DB 559 CATCGTTGGCCTCTTGGCTGCTGTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 618
QY 34 ---IleSerSerPro-CysPheIleGlyLeuThrGlyMetGlySerAlaThrGlnLeuAr 52
DB 619 CATTTACCTGCTCTCAATACCTGGGATCGCG---ATGGATCCACGTT----- 667
QY 52 gAlaArGsrLeuIleSerSerAlaValAlaThrAnSerLeuLHIsAPValGl 72
DB 668 -----GGCTTGACTCTCTTCGACATGGTTACCGGCTCGATATG----- 706
QY 72 yAlaThrValAlaValLeuGlYAlaATrAla----- 83
DB 707 -GGCACCCTGTCACATGTGACGCGGCTCGCTCGCTCGTGCAGATGAGATTAA 765
QY 84 -----LeuValIleuSerPheGluSerLeuThrLySArgAnValIleGlnGlnSe 100
DB 766 GCCTGTGAGGCAATGCGCAATGAAACCAACCACTATGCTTCATGATGAGGCGAC 825
QY 100 rLeuSerArgLySLeuValIleIleuSerGlyLeuPheValIleuAlaTrpPro1 120
DB 826 GATCACCCTGGGATCGTTCTAGCACTC---GGCATCATCTTTGGGATTCGCGGCGCAT 882
QY 120 ePheSerGlySerThrGlnAlaArg-----TyrPheAlaAlaPheVa 134
DB 883 GATGACTGATCTTCACCGCTACTGTTCCATCTGTGGGCAATTGGGCACTGTTGT 942
QY 134 1-----ProLeuValAsnGlyLeuAr 141
DB 943 CATCGTGGGTACGTTCTCTTCTACACAGCGCTTCATGCTGTGGGCGATGGCTGG 1002
QY 141 gLeuValIleAsnGlyLeuSerIleSerProAnSerMetLeuIleLySerValThrAr 161
DB 1003 CAAGTTATCGGTGACACCGTTGGGACGTTGAAAGCCTTGCTGGGACCACTCCAAAGC 1062
QY 161 gGluGlyArgAlaGluGluLeuLeuLySgLyProLeuPheTyrValIleuAlaLeuLeph 181
DB 1063 AAACCTCGCGCTGACCGCAACTACAGCGTTGCACTGACGTGGGCACTGCTTAGTG-- 1120
QY 181 eSerAlaValPhePheTrpArgGluSerProIleGlyMetIleSerLeuAlaMetMeCy 201
DB 1121 -----ACTGCAATTTGGCATGCTTTCTGCAACCATAGAG-- 1153
QY 201 eGlyGlyAspGlyIleAlaAspIleMetGlyArgLySArgLySerThrLyIleProTy 221
DB 1154 -----GACGCAAGTCCGACATGATGCGGAGACGATACCGGACGATTAATCTCGCA 1206
QY 221 rAnProArgLySserTrpAlaGlySerIleSerMet 233
DB 1207 GGGACCAACCAAC-----GGTTCATCAACCATG 1234

RESULT 6
US-09-328-352-3131
/ Sequence 3131, Application US/09328352
/ Patent No. 6562958
/ GENERAL INFORMATION:

```

```
/ APPLICANT: Gary L. Breton et al.
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
/ FILE REFERENCE: GTC99-03PA
/ CURRENT APPLICATION NUMBER: US/09/328,352
/ CURRENT FILING DATE: 1999-06-04
/ NUMBER OF SEQ ID NOS: 8252
/ SEQ ID NO 3131
/ LENGTH: 858
/ TYPE: DNA
/ ORGANISM: Acinetobacter baumannii
US-09-328-352-3131

Alignment Scores:
Pred. No.: 0.0115 Length: 858
Score: 100.50 Matches: 48
Percent Similarity: 38.9% Conservative: 40
Best Local Similarity: 21.2% Mismatches: 83
Query Match: 6.5% Indels: 55
Gaps: 8
DB: 3

US-10-634-548-2 (1-304) x US-09-328-352-3131 (1-858)

QY 70 AspValGlyAlaThrValAlaValLeuGlyAlaTyrAlaLeuValLeuSerPheGlu 89
DB 181 GAAGTTGGTGCAGTAGTAACCTTAAGCTGGGTTATGCTACTGCTGCTTTGTT 240
QY 90 Ser-----LeuThrIlyAaRghAnValIleGlnGlnSerLeuSerArgIlyLeuVal 106
DB 241 TCTGCGCTGCTCTTATCTTCTTCATGACATGCTTACTGTTATGCTGCTTCAATTGTA 300
QY 107 HisIleuSerGlyLeuLeuPheValLeuAlaTProIlePheSerGlySerThrGlu 126
DB 301 ACTGGCTGCTAGTGGTATTATGGGCAAGCTTTTCCGAATTCGATGCGGTGATTAAC 360
QY 127 AlaArgTyrPheAlaAlaPheValProLeuValAaGlyLeuArgLeuValIleAaGly 146
DB 361 GCGACCCATAT-----GTCATTGCTTATCTTATCTGTCAGCA 402
QY 147 LeuSerIleSerProAaSerMetLeuIleIySerValThrArgGluIyArgAlaGlu 166
DB 403 GTG----- 405
QY 167 GluLeuLeuIyGlyProLeuPheTyrValLeuAlaLeuPheSerAlaValPhe--- 185
DB 406 -----ACTGCAATCTTTGTA 420
QY 186 PheTyrArgIlySerProIleGlyMetIleSerLeuAlaMetMetCys---GlyGlyAaP 204
DB 421 GTATGCAAGCTCACCATGCTGATGATGTTGTTCTTCTGCTGCTGCTGCTGCTGCAAT 480
QY 205 GlyIleAlaSerPheIleMetGlyArgIlyPheGlySerThrIlyIle-----ProTyrAa 222
DB 481 AGTGAGCTTACTTGTAGGCGGTAAATTCGTAAGAAATTCGCGCTACAGTAAGT 540
QY 223 ProArgIlySerTyrPheValGlySerIleSerMetPheIlePheGlyPhePheIleSerIle 242
DB 541 CCTAATAATATCGTAAGT-----TTATATGCGGTATCTCACTCA 585
QY 243 AlaLeuLeuTyrTyrTyrSerSerLeuGlyTyrLeuHisMetAaSerTyrGluThrThrIle 262
DB 586 ATTATGTAATGCTGTGCTGATATATCATATTTAAACCTTAAGTGG-----GTT 636
QY 263 GlnArgValAlaMetValSerMetValAlaThrValValGlySerLeuProIleThrAaP 282
DB 637 CAACAACCTTATTTGATTTGCTTAACTACAGAGTTCGTTCAAGTATTAAGTAT 696
QY 283 GlnLeuAaPheAaIle 288
DB 697 TTATTTGAATCAATGATC 714

RESULT 7
US-09-540-236-198
```

```
/ Sequence 198, Application US/09540236
/ Patent No. 6673910
/ GENERAL INFORMATION:
/ APPLICANT: Gary L. Breton et al.
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATAR
/ FILE REFERENCE: 2709.2005-001
/ CURRENT APPLICATION NUMBER: US/09/540,236
/ CURRENT FILING DATE: 2000-04-04
/ NUMBER OF SEQ ID NOS: 3840
/ SEQ ID NO 198
/ LENGTH: 891
/ TYPE: DNA
/ ORGANISM: M. catarrhalis
US-09-540-236-198

Alignment Scores:
Pred. No.: 0.0122 Length: 891
Score: 100.50 Matches: 40
Percent Similarity: 45.9% Conservative: 28
Best Local Similarity: 27.0% Mismatches: 51
Query Match: 6.5% Indels: 29
Gaps: 9
DB: 3

US-10-634-548-2 (1-304) x US-09-540-236-198 (1-891)

QY 173 LeuPheTyrValLeuAlaLeuPheSerAlaValPhe-----PheTyr 187
DB 358 CTGGGTATATGCGGCGCTATCTCAACCGCTGATACCCGATGATGAGGCTTTGG 417
QY 188 ArgGlySerPro-----IleGlyMetIleSerLeuAlaMetMetCysGlyIyAaPgly 205
DB 418 CAGATGCGCTGCTGCTGCTATGATGATGTTTATGATGATGATGATGATGATGATGATG 477
QY 206 IleAlaSerPheMetGlyArgIlyPheGlySerThrIlyIleProTyrAaP-----Pro 223
DB 478 ---GCTTATTTTATGCGCAAAATTTGGCAAGCAAAATTTGGCAAGCAATGATACACC 534
QY 224 ArgIlySerTyrPheValGlySerIleSerMetPheIlePheGlyPhePheIleSerIleAla 243
DB 535 AATAAAGATGACGAAGCTTATGCGGTGATGATGATGATGATGATGATGATGATGATG 594
QY 244 LeuLeuTyrTyrTyrSerSerLeuGly-----TyrLeuHisMetAaP 258
DB 595 GTTGATATATTTGACACCTGCTGATGATGATGATGATGATGATGATGATGATGATG 654
QY 259 GluThrThrLeuGlnArgVal-----AlaMetValSerMetValAla 272
DB 655 CTAACTGATCTTCCAGATATCTGGGCACTTATTTGAATGATCTAAACGCGTGCA 714
QY 273 ThrValValGlySer-----LeuPro-----IleThrAaP 285
DB 715 GGTATTAAGATTTCTGGGCAATTTCTACAGATGATGCGGATTTTGAATGATGATGAT 774
QY 286 AaPheIleSer-----ValProLeu 292
DB 775 TCGCTATTTGACGATACGATTA 798

RESULT 8
US-09-596-002-37
/ Sequence 37, Application US/09596002
/ Patent No. 6632636
/ GENERAL INFORMATION:
/ APPLICANT: Lagace, Robert, E.
/ APPLICANT: Patterson, Chandra
/ APPLICANT: Berg, Kim, L.
/ TITLE OF INVENTION: NUCLEOTIDE SEQUENCES OF MORAXELLA CATARRHALIS GENOME
/ FILE REFERENCE: BM-0008-4-US
/ CURRENT APPLICATION NUMBER: US/09/596,002
/ CURRENT FILING DATE: 2000-06-16
/ PRIOR APPLICATION NUMBER: 60/140,121
/ PRIOR FILING DATE: 1999-06-18
/ NUMBER OF SEQ ID NOS: 41
US-09-540-236-198
```

```
! TYPE: DNA
```

: Sequence 964, Application US/09328352
: Patent No. 6562958

; Sequence 964, Application US/09328352
; Patent No. 6562958

```

; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 964
; LENGTH: 1548
; TYPE: DNA
; ORGANISM: Acinetobacter baumannii
US-09-328-352-964

Alignment Scores:
Pred. No.: 0.047 Length: 1548
Score: 99.00 Matches: 60
Percent Similarity: 41.04 Conservative: 40
Best Local Similarity: 24.64 Mismatches: 94
Query Match: 6.44 Indels: 50
DB: 3 Gaps: 14

US-10-634-548-2 (1-304) x US-09-328-352-964 (1-1548)
Qy 7 LeuSerProIleAsnHISGlnLeuCybArgPheGlyAsnAsnSerLeuThrThriHisArg 26
Db 67 CTCACCCCAATTAAAGTTAAATTAAGTAAAGTCTTCATCTGCATTAATGCAAGCGTG 126
Qy 27 PheCySerProGlyPheLeuIleSerSerProCybPheIleGlyLeu---ThriGlyMet 45
Db 127 GCTGTGTAGAGCAAAATGTAAATGAAGACGTAACGCTTACGGTATTAATACAGCTTTT 186
Qy 46 Gly-----SerAlaThrGlnLeu-----ArgAlaArgArgSerLeu 57
Db 187 GCGTACTGCTTCAACTAAATTCGCGCTGAAAGATTAGAAAATTACAGCGTTCATTA 246
Qy 58 Ile---SerSerAlaValAlaIleThrAsnSerLeuLeuHisAspValGlyAlaThrValAla 76
Db 247 GTCTTTCGCATCGCGCAGGTGTAGGCAAGACCTTGATGATGCAATGCTTCGTTAATT 306
Qy 77 ValLeuGlyGlyAlaIleValAlaLeuValLeuSerPheGlnSerLeuThrIleArgAsnVal 96
Db 307 ATTCTATTGAAGCAAAATAGCTTGACAGCTGCTTCGCTGAT---CGTCGTAAAGTC 363
Qy 97 IleGlnGlnSerLeu-----SerArgIleValAlaHisIle----- 108
Db 364 ATTGATGCTTGTGGCCTTAATTAATGCTGAAGTTTACCAACACATTCCACTTAAAGCT 423
Qy 109 -----LeuSerGlyLeuLeuPheValLeuAlaIlePro-----IlePheSerGly 123
Db 424 TCAGTAGGGGCTTCTGGTGACTTGGCACCCCTTGACACACATGCTTTGGTITTTACTTGT 483
Qy 124 SerThrGlnAlaArgIlePheAlaIlePheValProLeuValAsnGlyLeuArgLeuVal 143
Db 484 GAAAGTAAAGCAGCTTAATGAAGTGAATGTTACACAGACGTTGAAGCTTTAAATTCGA 543
Qy 144 IleAsnGlyLeuSerIleSerProAsnSerMetLeuIleIleSerValIleThrArgGly 163
Db 544 -----GGTTTGAACCAATTTCTTAAAGCGGCTTAA-----GAAAGCT 579
Qy 164 ArgAlaGlnGlnLeuLeuLeuGly-----ProLeuPheTyValLeuAlaLeuLeu 180
Db 580 TTGGCA-----CTTTTAAATGTTACGACAGTTTGACAGCTTACGCTTTACGTGGCTTA 633
Qy 181 PheSerAlaValPhePheTrpArgIleSerProIleGlyMetIleSerLeuAlaMetMet 200
Db 634 TTTGAAGACAGA-----GATTATTTGCTGCGCGACAGATT 669
Qy 201 CysGlyGlyAspGlyIleAlaAspIleMetGlyArgIlePheGlySerThrIleIlePro 220
Db 670 TGTGTGTGCTTAATGTTGAAGCAATGCTGTG-----TCACGTGCACCT 714
Qy 221 TyrAsnProArg 224

```

```

Db 715 TTGCATGGCGCT 726
RESULT 11
US-09-710-279-2167
; Sequence 2167, Application US/09710279
; Patent No. 6703492
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/09/710,279
; CURRENT FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: 60/164,258
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 2167
; LENGTH: 1152
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
US-09-710-279-2167

Alignment Scores:
Pred. No.: 0.0334 Length: 1152
Score: 98.50 Matches: 70
Percent Similarity: 40.94 Conservative: 58
Best Local Similarity: 22.44 Mismatches: 115
Query Match: 6.44 Indels: 71
DB: 3 Gaps: 13

US-10-634-548-2 (1-304) x US-09-710-279-2167 (1-1152)
Qy 13 GlnLeuCybArgPheGlyAsnAsnSerLeuThrThriHisArgPheCySerProGlyPhe 32
Db 181 CAATTCATGCTTTTAAATCAAGTACGTCGCAAAATATTCATTTGAACA----- 231
Qy 33 LeuIleSerSerProCybPheIleGlyLeuThrGlyMetGlySerAlaThrGlnLeuArg 52
Db 232 ATCATTTAGTATTCGA-----TGGTGACTGTTTAAAGTCAAGCATATTTATTTTG 282
Qy 53 AlaArgArgSerLeuIleSerSerAlaValAlaIleThrAsnSerLeuLeuHisAspValGly 72
Db 283 ATTAACAATGTAATTTATGTTAGTGTGTAAGT-----TTAGAAAAGCTTCT 336
Qy 73 AlaThrValAlaValAlaLeuGlyAlaIleValAlaLeuValLeuSerPheGlnSerLeuThr 92
Db 337 AAAATTAATGATGCTCTAATTTATTTATCTTTTAAATATGTTGACACAAATCTTTTACT 396
Qy 93 LysArgAsnValIleGln-----GlnSerLeuSer 102
Db 397 TTAGAAGTGCTTTAGAAAGGTGATGATATATCTGCAACTCGAGTTGAAGATATGCT 456
Qy 103 ArgIleLeuValHisIleLeuSerGlyLeuLeuPheValLeuAlaIleProIlePheSer 122
Db 457 -----ATTCAGAGTGTTACTATTTGGTTTGAACAAATCGTTTAAAG 498
Qy 123 -----GlySerThrGlnAlaArgIlePheAlaIlePheValPro----- 135
Db 499 CTGTCCTAGGTATCAACCGCAATGATTAATCTTATGACAGCTTAAATAATATGACG 558
Qy 136 -----LeuValAsnGlyLeuArgLeuValIleAsnGly 146
Db 559 ATAAAGCTTCAGACATTTCAATTGCGTATGATGATATTTTAAATTCGCTGTGGCTGGA 618
Qy 147 LeuSerIleSerProAsnSerMetLeuIleIleSerValIleThrArgGlnGlyArgAlaGln 166
Db 619 TTACCTATATTTCTTGGCGCTTAAACAATTTGGTTTACCAACCCCAAGAGG----- 669
Qy 167 GlnLeuLeuGlyGlyProLeuPheTyValLeuAlaLeuLeuPheSerAlaValPhePhe 186

```

```

Db      670 -----CCTGGCTTATTTATTTAAAGTTTACCACTAGATTTAGCGAATGAC-TTT 719
Qy      187 TTPARGUSERPROILEGLYMETILSERLEUAMETCYGGLYASP-GLY11 206
Db      720 TGGTACATCTTTTAACTTTATTTTACTATTTATTTTGGGCAATTAACTCTTC 779
Qy      206 eAlaSerMetGlyArgLysPheGlySer---ThrLysIleProTyrAsnProArgL 225
Db      780 TATATCATTTATTAAGTAAATGTCATATTTTCTAAATAATGATAATAGTAAAGCA 839
Qy      225 sSerTrpAla-----GlySerIleSerMetPheIlePheGlyPhePheIleSerIleAl 243
Db      840 AAAAGTGCACATCATAGAGTACTGTAATTTATTC-----ATTAGTATCCC 887
Qy      243 aleuLeuYrTyTrpSerSerLeuGlyTyTrpLeuHisMetLeuTrpGluThrLeuGl 263
Db      888 AGCAACATTATCTTTTAAAGTCAATTCATTTGCGTTTGGC----- 930
Qy      263 nArgValAlaMetValSerMetValAlaThrValAlaGluSerLeuProIleThrAspGl 283
Db      931 -----GCTGTCACATATTGATAATAG-----GATTT 959
Qy      283 nLeuAspAspAsnIleSerValProLeuAlaThrIle 295
Db      960 TATTGTATCTAATATTCTTATGCAATTAGGGGCACCTA 996

RESULT 12
US-09-134-001C-1098
; Sequence 1098, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 1098
; LENGTH: 1359
; TYPE: DNA
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-1098

Alignment Scores:
Pred. No.: 0.0439      Length: 1359
Score: 98.50      Matches: 70
Percent Similarity: 40.9%      Conservative: 58
Best Local Similarity: 22.4%      Mismatches: 115
Query Match: 6.4%      Indels: 71
DB: 3      Gaps: 13

US-10-634-548-2 (1-304) x US-09-134-001C-1098 (1-1359)
Qy      13 GlnLeuCyArgPheGlyAsnAsnSerLeuThrThiHisArgPheCySerProGlyPhe 32
Db      388 CAATCATGCTTTTAAATCAAGTACGCTGACAAATATTCATTTGAACA----- 438
Qy      33 LeuIleSerSerProCySerPheIleGlyLeuThrGlyMetGlySerAlaThrGlnLeuArg 52
Db      439 ATCATTTAGTAATCCA-----TGGTACCTGCTTTTACGTCACAGGCAATTTATTG 489
Qy      53 AlaArgArgSerLeuIleSerSerAlaValAlaThrAsnSerLeuLeuHisAspValGly 72
Db      490 ATTAACAATGTAATCTTATGTTAGCTGTTGAAAAAGG-----TTAGAAAAAGCTTCT 543
Qy      73 AlaThrValAlaValLeuGlyGlyAlaTyAlaLeuValLeuSerPheGluSerLeuThr 92

```

```

Db      544 AAAATTAATGATGCTTATTAATTTATCTTTTAAATATTCGTTGACCAATCTTTACT 603
Qy      93 LysArgAsnValIleGln-----GlnSerLeuSer 102
Db      604 TTAAAGTGTCTTTAGAAAGTGTCATTAATTAATGCAACCTGAGTTGAAGATATGCT 663
Qy      103 ArgLysLeuValHisIleLeuSerGlyLeuLeuPheValLeuAlaTrpProIlePheSer 122
Db      664 -----ATTCAGAGGTGACTATTTTCGTTAGGGCAATCGTTTATACG 705
Qy      123 -----GlySerThrGluAlaArgTyPheAlaIlePheValPro----- 135
Db      706 CTGTCCTAGTACACCGGATGATTAATTAATTTTAAATTTTGTGCTGGCTGCA 825
Qy      136 -----LeuValAsnGlyLeuArgLeuValIleAsnGly 146
Db      766 ATTAAGCTTCAGACATTTCAATTTGTGTATGAAATTTTAAATTTTGTGCTGGCTGCA 825
Qy      147 LeuSerIleSerProAsnSerMetLeuIleLysSerValThrArgGluGlyArgAlaGlu 166
Db      826 TTACCTATATTCTCTGCGCTTAAACATTTGTTACCAACCCCAAGAGGC----- 876
Qy      167 GluLeuLeuYsGlyProLeuPheTyValLeuAlaLeuLeuPheSerAlaValPhePhe 186
Db      877 -----CCTGCTTATTTAATTTAAGTCTTACCACTAGATTTTACGAAATGAC-TTT 926
Qy      187 TTPARGUSERPROILEGLYMETILSERLEUAMETCYGGLYASP-GLY11 206
Db      927 TGGTACATCTCTTTAACTTTATTTTAACTTATTTTAACTTATTTTGGGCAATGAC-TTT 986
Qy      206 eAlaSerMetGlyArgLysPheGlySer---ThrLysIleProTyrAsnProArgL 225
Db      987 TATATCATTTATTAAGTAAATGTCATCTAATTTTACTAAATAATGATAATAGTAAAGCA 1046
Qy      225 sSerTrpAla-----GlySerIleSerMetPheIlePheGlyPhePheIleSerIleAl 243
Db      1047 AAAAGTGCACATCATAGAGTACTGTAATTTATTC-----ATTAGTATCCC 1094
Qy      243 aleuLeuYrTyTrpSerSerLeuGlyTyTrpLeuHisMetLeuTrpGluThrLeuGl 263
Db      1095 AGCAACATTATCTTTTAAAGTCAATTTGCGTTTGGC----- 1137
Qy      263 nArgValAlaMetValSerMetValAlaThrValAlaGluSerLeuProIleThrAspGl 283
Db      1138 -----GCTGTCACATATTGATAATAG-----GATTT 1166
Qy      283 nLeuAspAspAsnIleSerValProLeuAlaThrIle 295
Db      1167 TATTGTATCTAATATTCTTATGCAATTAGGGGCACCTA 1203

RESULT 13
US-09-710-279-3427/C
; Sequence 3427, Application US/09710279
; Patent No. 6703492
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/09/710,279
; PRIOR FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: 60/164,258
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 3427
; LENGTH: 3108
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: nucleic acid sequence
US-09-710-279-3427

```

Alignment Scores:

Pred. No.: 0.171 Length: 3108
 Score: 98.50 Matches: 70
 Percent Similarity: 40.98 Conservative: 58
 Best Local Similarity: 22.44 Mismatches: 115
 Query Match: 6.4% Indels: 71
 Gaps: 13

US-10-634-548-2 (1-304) x US-09-710-279-3427 (1-3108)

QY 13 GlnLeuCyArgPheGlyAsnAsnSerLeuThrHsIArgPheCySerProGlyPhe 32
 Db 3087 CAATCATGCTTTTAAATCAAGTCCGACCAATATTCATTTGAACA----- 3037
 QY 33 LeuIleSerSerProCyPheIleGlyLeuThrGlyMetGlySerAlaThrGlnLeuArg 52
 Db 3036 ATCATTAAGTAAATCC-----TGCTTGAAGTCTTTAGGCAAGCATTTATTTTG 2986
 QY 53 AlaArgArgSerLeuIleSerSerAlaValAlaThrAsnSerLeuLeuHisAspValGly 72
 Db 2985 ATTAACATGCTTAATGTTATGTTAGCTGTGAAAAAGCT-----TTAGAAAAAGCTTCT 2932
 QY 73 AlaThrValAlaValLeuGlyGlyAlaTyralLeuValLeuSerPheGlnSerLeuThr 92
 Db 2931 AAAATTAATGATGCCCTCTATTTATTTATTTTAAATTAATCGTTGACCAATCTTTAACT 2872
 QY 93 LysArgAsnValIleGln-----GlnSerLeuSer 102
 Db 2871 TTAGAGGCTGTTTGAAGGTGTGACGTATATACGCAACCTCGAGTTGAAGATATGCT 2812
 QY 103 ArgLysLeuValHisIleLeuSerGlyLeuLeuPheValLeuAlaTrpProIlePheSer 122
 Db 2811 -----ATTCAGGTGATGACTATTTGCGTTAGCAATCGTTTATACG 2770
 QY 123 -----GlySerThrGlnAlaArgTyrrPheAlaAlaPheValPro----- 135
 Db 2769 CTGTCCCTAAGGACCAACCGAATGATTAATGACAGTATGACCTTAATAAATATGACG 2710
 QY 136 -----LeuValaenglyLeuArgLeuValIleasnGly 146
 Db 2709 ATAAAGTCTTGACGACTTTCATTTGCGTAATGAATATTTAAATTTCTGCTGTGCTCGA 2650
 QY 147 LeuSerIleSerProAsnSerMetLeuIleYserSerValThrArgGlnGlyArgAlaGln 166
 Db 2649 TTAGCTATATTTCCCTGCGCTTAAACATTTGGTTACCAACCCCAAGAAAGGC----- 2599
 QY 167 GluLeuLeuGlyProLeuPheTyrrValLeuAlaLeuLeuPheSerAlaValPhePhe 186
 Db 2598 -----CCGGCTTATTTATTTAAAGTTTACACACTAGTATTTAGGAATATGAC-TTT 2549
 QY 187 TrpArgGlnSerProIleGlyMetIleSerLeuAlaMetMetCyGlyGlyAsp-GlyIle 206
 Db 2548 TGTGACATCTTTTACCTTATTTTATTTTACTTATTTATTTGCGGCAATTAACCTCTTC 2489
 QY 206 AlaAspIleMetGlyArgLysPheGlySer---ThrLysIleProTyrrAsnProArgLys 225
 Db 2488 TATATCATTAATTAAGCTTAAATGATTAATTTTCTAAATAAATGATAATAGAAAGCA 2429
 QY 225 sSerTrpAla-----GlySerIleSerMetPheIlePheGlyPhePheIleSerIleAl 243
 Db 2428 AAAAGTGCATTCATAGGATGATATCTGTTATTTATC-----ATTAGATATCC 2381
 QY 243 AlaLeuLeuTyrrTyrrSerSerLeuGlyTyrrLeuHisMetAsnTrpGlnThrLeuGln 263
 Db 2380 AGCAACATTAATCTTTAGTACGCTAAGCATTTGGGTTTGCC----- 2338
 QY 263 nArgValAlaMetValSerMetValAlaThrValValGlnSerLeuProIleThrAspGln 283
 Db 2337 -----GCTGTACATATTTATTAATATG-----GATTT 2309
 QY 283 nLeuAspAsnAsnIleSerValProLeuAlaThrIle 295
 Db 2308 TATGTATCATTAATTTCTTATGCGCATTTAGGGGCACTA 2272

RESULT 14

US-09-543-681A-282

/ Sequence 282, Application US/09543681A

/ Patent No. 6605709

/ GENERAL INFORMATION:

/ APPLICANT: GARY BRETON

/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS

/ FILE REFERENCE: 2709.1002-001

/ CURRENT FILING DATE: 2000-04-05

/ PRIOR APPLICATION NUMBER: US 60/128,706

/ NUMBER OF SEQ ID NOS: 8344

/ SEQ ID NO 282

/ LENGTH: 903

/ TYPE: DNA

/ ORGANISM: Proteus mirabilis

US-09-543-681A-282

Alignment Scores:

Pred. No.: 0.03 Length: 903
 Score: 97.50 Matches: 66
 Percent Similarity: 39.2% Conservative: 38
 Best Local Similarity: 24.9% Mismatches: 78
 Query Match: 6.3% Indels: 83
 Gaps: 17

US-10-634-548-2 (1-304) x US-09-543-681A-282 (1-903)

QY 31 GlyPheLeuIleSerSerProCyPheIleGlyLeuThrGlyMetGlySerAlaThrGln 50
 Db 214 GGTATTTTACTCTCC-----AGATCCATATGGA-----ACCCCA 249
 QY 51 LeuArgAlaArgArgSerLeuIleSerSerAlaValAlaThrAsnSerLeuLeuHisAsp 70
 Db 250 ATGCGTGCCTTAATCTTTATTC-----CGTTTGGCTATCATTAATAGTAATGA 303
 QY 71 Val-----GlyAlaThrValAlaValLeuGlyGlyAlaTyralAlaLeuValLeuSerPhe 88
 Db 304 ATTGCAAGTGTGACCACTCTTGTCTTGAAGTGTGCTTAATATTTGGCTAAT----- 354
 QY 89 GluSerLeuThrLysArgAsnValIleGlnIleSerLeuSerArgLysLeuValHisIle 108
 Db 355 AGTGTGTTAAATTAATAAGCCGCTGCAATTAAGTAATGTGCTAGTGTGTCATGATA 414
 QY 109 LeuSerGlyLeuLeuPheValLeuAla----- 117
 Db 415 CTTGCA---GTGCTATTTATTAACGGCTATTTTCGCGAGTTTATCAATTAATGATAGTACT 471
 QY 118 ---TrpProIlePheSerGlySerThrGlnAlaArgTyrr---PheAlaIlePheVal--- 134
 Db 472 ACTTGG-----TAAATGCTATTAACACAGTTAACTTCGATCAACGCAATTTATGCG 525
 QY 135 ---ProLeuValaenglyLeuArgLeuValIleasnGlyLeuSerIleSerProAsnSer 153
 Db 526 GGCCTATTTTAAAGCGGCACTGTAAATGCGATTTGCGGTTTAACTTAAATGATTAATGT 585
 QY 154 MetLeuIleYserSerValThrArgGlnGlyArgAlaGlnGluLeuLeuGlyProLeu 173
 Db 586 GCATTA-----CCGCTA 597
 QY 174 PheTyrrValLeuAlaLeuLeuPheSerAlaValPhePheTyrrArgGlnSer---ProIle 192
 Db 598 TTAAGTGCATCGCATTAATATATGATGCTTATTTGCTTACCTCAAGCTTTGAATTTG 657
 QY 193 GlyMetIleSerLeuAlaMetMetCyGlyGlyArgGlyIleAlaAsp-----Ile 209
 Db 658 GGTTCATTAACCTCGTACCAAAAAGCGGTGATTAATTAATTAATTAATTAATTAATTA 717
 QY 210 MetGlyArgLysPheGlySerThrLysIle----- 219

```

Db      718 ATGGGATTAAATGATGGCCCTAACATTAGGCTTGATGATAGACACCGTTATTA 777
Qy      220 ---ProTyrAnProArgLysSerTrpAlaGlySerIleSerMetPheIlePhe 238
Db      778 CGCAAAATAATATCC-----TCGTGTTCTTTAATAGTGTGTAGTTT 819
Qy      239 -----PheIleSerIleAlaLeuLeuTyrTyrTyrSerSerLeu 251
Db      820 ATTTTGCTGTTTGTGTCATTTATTTGTCGTGCGCATTTCTA----- 864
Qy      252 GlyTyrLeuHisMet 256
Db      865 ---GGCTTACATATG 876

RESULT 15
US-09-543-681A-650
; Sequence 650, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 650
; LENGTH: 1227
; TYPE: DNA
; ORGANISM: Proteus mirabilis
US-09-543-681A-650

Alignment Scores:
Pred. No.: 0.0575      Length: 1227
Score: 97.00      Matches: 59
Percent Similarity: 36.5%      Conservative: 54
Best Local Similarity: 19.0%      Mismatches: 106
Query Match: 6.3%      Indels: 91
DB: 3      Gaps: 11

US-10-634-548-2 (1-304) x US-09-543-681A-650 (1-1227)
Qy      10 ILeaenHisGlnLeuCyArgPheGlyAsnAsnSerLeuThr-----HisArg 26
Db      300 TTGGGTAGACAGTGTGCGGTTTACTAACAGCGCTTTCTTACAAAGCTTTGGCGTCTG 359
Qy      27 PheCySerProGlyPheLeuIleSerSerProCyPheIleGlyLeuThrGly-MetG1 46
Db      360 CTCCTCAGCCGT-----TATTTGGCAAGCAGCTAGTATAGA 395
Qy      46 ySerAlaThrGlnLeuArgAlaArg------SerLeuIleSerSerAlaValAl 63
Db      396 CCGTTTGTAGTGTGCGGCTCAGCACAACGTTTTCGCATCATATTATGCCATTGGTCCCT 455
Qy      63 aThrAsnSerLeuLeuHisAspValGlyAlaThrValAlaValLeuGlyGly----- 80
Db      456 TTCCCGAGCTTAGCGGCACTCATTTGCTCTTTCTGTAAACACAGGTGGTGGCGCAT 515
Qy      81 -----AlaTyrAlaLeuValIleuSerPheGluSerLeuThrLysAr 94
Db      516 GATCTTTATGTTGCTGATGCAATTTCTTGTATTATATGCTAACCCACATTTATGCTAA 575
Qy      94 gAnValIleGlnGlnSerLeuSerArgLysLeuValHisIleLeuSerGlyLeuLeuPh 114
Db      576 AAATATTAAACAAATCAGTTATGTATGATCAGGCCCATCTACAGTATCA---TTTTAAC 632
Qy      114 eValIleuAlaTyrProIlePheSerGlySer-----ThrGluAlaArgTyrPheAlaAl 132
Db      633 TTTCCTTAAGTACCTATATCTATAGTCCGGTAAAGCATATTATGCTTACGCTCTGATGTCGGG 692
Qy      132 apHeValProLeuValAsnGlyLeuArgLeuValIleAsnGlyLeuSerIleSerProAs 152

```

```

Db      693 TTTCTTGCTTGTTAGACAGGCTCTCCGATTTATCTTGAGAAGATGGCTTATTTCACACA 752
Qy      152 nSerMetLeuIleLysSerValThrArgGlnGlyArgAlaGluGluLeuLysGlyPr 172
Db      753 AGATATCGGCTTAATGCTATATTCCAA----- 780
Qy      172 OleuPheTyrValLeuAlaLeuLeuPheSerAlaValPhePheTrpArgGluSerProI 192
Db      780 ----- 780
Qy      192 eGlyMetIleSerLeuAlaMetMetCysGlyGlyAspGlyTyrIleAlaAspIleMetGlyAr 212
Db      781 -----ACCATCGCATTTATGTTAGGTGTATGTTGTGCGCATTTATTTACAA 830
Qy      212 GlyPheGlySerThrLysIleProTyrAsnProArgLysSerTrpAlaGlySerIleSe 232
Db      831 AATGTCGGGTAAACACTGTATCCT----- 855
Qy      232 rMetPheIlePheGlyPhePheIleSerIleAlaLeuLeuTyrTyrTyrSerSerLeuG1 252
Db      856 -TTATGTATTATGCTTAGCGAGTACGATGACAGCGTTGTTCTTATGTCGCAAA----- 909
Qy      252 YTyrLeuHisMetAsnTrpGluThrThrLeuGlnArgValAlaMetValSerMetValAl 272
Db      910 -TATACACACCCCAATTAAACAAATTTTTCATTTGTATTATGCTGATGATG-- 966
Qy      272 aThrValAlaLysSerLeuProIleThrAspGlnLeuAspAsnIleSerValProLe 292
Db      967 -----AACGGTCTTCTTATCCAT 986
Qy      292 uAla-----ThrIleuAlaAlaTyr 299
Db      987 TGCTGTCTTATGCTATCAGCCTAC 1014

```

Search completed: March 16, 2006, 23:47:52
Job time : 1016 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 BioCeleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 16, 2006, 23:30:08 ; Search time 904 Seconds
(without alignments)
2780.853 Million cell updates/sec

Title: US-10-634-548-2

Perfect score: 1536
Sequence: 1 MAATLPLSPINHCRCFGNN.....DDNISVPLATIAATLALSRGY 304

Scoring table:

BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODL=frame+g2n.model -DEV=x1h
-Q=/abse/ABSSMBR.epool/US10634548/runat.16032006.084440.10458/app.query.fasta_1
-DB=Published Applications NA Main -QFMT=fastap -SUFFIX=p2n.rnpbm
-MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1
-MATRIX=blosum62 -TRANS=human40.cdi -LIST=45 -DOCALLIGN=200 -THR_SCORE=pct
-THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pco -NORM=ext
-HEASize=500 -MINLEN=2000000000 -HOST=abse02h
-USER=US10634548 @CGN 1 1 1026 @runat.16032006.084440.10458 -NCPU=6 -ICPU=3
-NO MAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database: Published Applications NA Main:

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
6: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1536	100.0	1015	US-10-425-114-21028	Sequence 21028, A
2	1536	100.0	1091	US-10-634-548-1	Sequence 1, Appli
3	1523	99.2	980	US-10-425-114-14606	Sequence 14606, A
4	1284.5	83.6	1897	US-10-634-548-3	Sequence 3, Appli
5	1266	82.4	1062	US-10-425-114-29129	Sequence 29129, A
6	870	56.6	1007	US-10-425-114-20410	Sequence 20410, A
7	794	51.7	1085	US-10-425-114-22326	Sequence 22326, A

8	794	51.7	1974	7	US-10-437-963-77300	Sequence 77300, A
9	794	51.7	2514	7 <td>US-10-437-963-77301</td> <td>Sequence 77301, A</td>	US-10-437-963-77301	Sequence 77301, A
10	788	51.3	1033	7 <td>US-10-425-114-28593</td> <td>Sequence 28593, A</td>	US-10-425-114-28593	Sequence 28593, A
11	777	50.6	1123	8 <td>US-10-425-114-118753</td> <td>Sequence 118753, A</td>	US-10-425-114-118753	Sequence 118753, A
12	776	50.5	798	7 <td>US-10-425-114-23255</td> <td>Sequence 23255, A</td>	US-10-425-114-23255	Sequence 23255, A
13	773	50.3	976	7 <td>US-10-425-114-26551</td> <td>Sequence 26551, A</td>	US-10-425-114-26551	Sequence 26551, A
14	751	48.9	723	7 <td>US-10-425-114-25671</td> <td>Sequence 25671, A</td>	US-10-425-114-25671	Sequence 25671, A
15	751	48.9	753	7 <td>US-10-425-114-26498</td> <td>Sequence 26498, A</td>	US-10-425-114-26498	Sequence 26498, A
16	742	48.3	937	7 <td>US-10-425-114-26454</td> <td>Sequence 26454, A</td>	US-10-425-114-26454	Sequence 26454, A
17	646	42.1	674	7 <td>US-10-425-114-16877</td> <td>Sequence 16877, A</td>	US-10-425-114-16877	Sequence 16877, A
18	594	38.7	992	7 <td>US-10-424-599-112249</td> <td>Sequence 112249, A</td>	US-10-424-599-112249	Sequence 112249, A
19	574.5	37.4	1286	8 <td>US-10-425-115-172503</td> <td>Sequence 172503, A</td>	US-10-425-115-172503	Sequence 172503, A
20	573	37.3	916	8 <td>US-10-739-930-1001</td> <td>Sequence 3001, Ap</td>	US-10-739-930-1001	Sequence 3001, Ap
21	572.5	37.3	1225	7 <td>US-10-425-114-26462</td> <td>Sequence 26462, A</td>	US-10-425-114-26462	Sequence 26462, A
22	567.5	36.9	688	7 <td>US-10-425-114-22387</td> <td>Sequence 22387, A</td>	US-10-425-114-22387	Sequence 22387, A
23	567.5	36.9	907	7 <td>US-10-767-701-11210</td> <td>Sequence 11210, A</td>	US-10-767-701-11210	Sequence 11210, A
24	566.5	36.9	666	7 <td>US-10-425-114-4925</td> <td>Sequence 4925, Ap</td>	US-10-425-114-4925	Sequence 4925, Ap
25	566.5	36.9	955	8 <td>US-10-425-115-118754</td> <td>Sequence 118754, A</td>	US-10-425-115-118754	Sequence 118754, A
26	552.5	36.0	1091	7 <td>US-10-425-114-26298</td> <td>Sequence 26298, A</td>	US-10-425-114-26298	Sequence 26298, A
27	545.5	35.5	1431	8 <td>US-10-739-930-4866</td> <td>Sequence 4866, Ap</td>	US-10-739-930-4866	Sequence 4866, Ap
28	544.5	35.4	1430	8 <td>US-10-424-599-11025</td> <td>Sequence 71025, A</td>	US-10-424-599-11025	Sequence 71025, A
29	540.5	35.2	795	7 <td>US-10-425-114-13412</td> <td>Sequence 13412, A</td>	US-10-425-114-13412	Sequence 13412, A
30	539	35.1	867	7 <td>US-10-425-114-25373</td> <td>Sequence 25373, A</td>	US-10-425-114-25373	Sequence 25373, A
31	536.5	34.9	1046	7 <td>US-10-425-114-19295</td> <td>Sequence 19295, A</td>	US-10-425-114-19295	Sequence 19295, A
32	530.5	34.5	1162	7 <td>US-10-634-548-5</td> <td>Sequence 5, Appli</td>	US-10-634-548-5	Sequence 5, Appli
33	529	34.4	1057	7 <td>US-10-767-701-11766</td> <td>Sequence 11766, A</td>	US-10-767-701-11766	Sequence 11766, A
34	517	33.7	876	7 <td>US-10-437-963-97687</td> <td>Sequence 97687, A</td>	US-10-437-963-97687	Sequence 97687, A
35	517	33.7	1193	8 <td>US-10-739-930-1297</td> <td>Sequence 1297, Ap</td>	US-10-739-930-1297	Sequence 1297, Ap
36	517	33.7	1281	7 <td>US-10-767-701-11596</td> <td>Sequence 11596, A</td>	US-10-767-701-11596	Sequence 11596, A
37	512	33.3	1029	7 <td>US-10-425-114-14899</td> <td>Sequence 14899, A</td>	US-10-425-114-14899	Sequence 14899, A
38	511.5	33.3	648	7 <td>US-10-425-114-15929</td> <td>Sequence 15929, A</td>	US-10-425-114-15929	Sequence 15929, A
39	509	33.1	944	7 <td>US-10-425-114-19979</td> <td>Sequence 19979, A</td>	US-10-425-114-19979	Sequence 19979, A
40	509	33.1	1106	7 <td>US-10-425-114-28780</td> <td>Sequence 28780, A</td>	US-10-425-114-28780	Sequence 28780, A
41	508	33.1	948	7 <td>US-10-425-114-19863</td> <td>Sequence 19863, A</td>	US-10-425-114-19863	Sequence 19863, A
42	500.5	32.6	818	7 <td>US-10-437-963-13980</td> <td>Sequence 13980, A</td>	US-10-437-963-13980	Sequence 13980, A
43	490	31.9	1044	7 <td>US-10-425-114-26552</td> <td>Sequence 26552, A</td>	US-10-425-114-26552	Sequence 26552, A
44	486.5	31.7	1548	8 <td>US-10-425-115-44228</td> <td>Sequence 54228, A</td>	US-10-425-115-44228	Sequence 54228, A
45	483	31.4	1086	7 <td>US-10-767-701-11767</td> <td>Sequence 11767, A</td>	US-10-767-701-11767	Sequence 11767, A

ALIGNMENTS

RESULT 1
US-10-425-114-21028
; Sequence 21028, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jindong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(5313)B
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 21028
; LENGTH: 1015
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana columbia
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB378-077-D5_FLI
US-10-425-114-21028
Alignment Scores:
Pred. No.: 1.66e-177
Score: 1536.00
Percent Similarity: 100.0%
Best Local Similarity: 100.0%
Query Match: 100.0%
Length: 1015
Matches: 304
Conservative: 0
Mismatch: 0
Indels: 0

DB: 7 Gaps: 0
US-10-634-548-2 (1-304) x US-10-425-114-21028 (1-1015)
QY 1 MetAlaAlaThrLeuProLeuSerProIleAenHISGInLeuCyAArgPheGlyValAsn 20
DB 35 ATGGACGAACTTACTCTATCTCCGATCATCATAGTGTGTGCGGTCCGGAAAC 94
QY 21 SerLeuThrThrHISArgPheCysSerProGlyPheLeuIleSerSerProCysPheIle 40
DB 95 TCTTGGACGACTCACCGGCTTCTGTCTCTGCGCTTCTGTGATTCCTCTCTGTTCA 154
QY 41 GlyLeuThrGlyMetGlySerAlaThrGInLeuArgAlaArgArgSerLeuIleSerSer 60
DB 155 GGTGGACCGAAATGGGCTCTGCTACTAGTACGTGCTCGCTGCTTCTGATCTTCA 214
QY 61 AlaValAlaThrAsnSerLeuLeuHISAspValGlyAlaThrValAlaValLeuGlyGly 80
DB 215 GCAGTTGCGAAGAAATTCGCTGTGATGACGTCGAGCCACCGTGGCAGTGTGTGA 274
QY 81 AlaTyrAlaLeuValLeuSerPheGlyLeuThrIleValArgAsnValIleGInGInSer 100
DB 275 GCATACGGCGCTGTCTTAAGCTTCAGAGCTTCACCAAGCGAAAGTCATTCACAGAGT 334
QY 101 LeuSerArgGlyLeuValHISIleLeuSerGlyLeuLeuPheValLeuAlaTrpProIle 120
DB 335 TTGACGAGAAAGCTGTGATATCTCTCAGGCTGCTTCTGTTGCTGCTGCGCAATC 394
QY 121 PheSerGlySerThrGInAlaArgTyrPheAlaAlaPheValProLeuValLeuGlyLeu 140
DB 395 TTCACGCGATCGACCGAGGCTCGATATCTTGGCTTGTTCCTGTAGTGAATGGCTTA 454
QY 141 ArgLeuValIleAsnGlyLeuSerIleSerProAsnSerMetLeuIleYSerValThr 160
DB 455 AGCGTGTATTAACGAGCTATCCATTCCTCCAAATTCATCTAATCAATCCGTACA 514
QY 161 ArgGInGlyArgAlaGInGInLeuLeuGlyGlyProLeuPheTyrValLeuAlaLeuLeu 180
DB 515 AAGAGAGGAGAGAGAGAGAGAGTTCCTTAAGGCTTCTTCTTCAAGCTTCTTCTT 574
QY 181 PheSerAlaValPhePheTrpArgIleUserProIleGlyMetCIIeSerLeuAlaMetMet 200
DB 575 TCTCTGCGGCTTCTTCTGAGAGAGTCTCTATCGATATCGTATCGTATCGCAATGAG 634
QY 201 CysGlyGlyAspGlyTyrIleAlaAspIleMetGlyArgGlyPheGlySerThrIlePro 220
DB 635 TGTGTGGCGAGTGAATAGCTGATATATAGGAGCTTGGGTCATTAAGTACTT 694
QY 221 TyrAsnProArgIleSerTrpAlaGlySerIleSerMetPheIlePheGlyPhePheIle 240
DB 695 TACAACCCAGAGAGAGAGTGGGAGAGAGATCTCCATGTTCACTTGGGCTTCTTCA 754
QY 241 SerIleAlaLeuLeuTyrTyrTyrSerSerLeuGlyTyrLeuHISMetAsnTrpGInThr 260
DB 755 TCCATCGACTTACTTACTTACTTACTCAAGCTTGGGTACTTCACTGAACTGGGAAAG 814
QY 261 ThrLeuGInArgValAlaMetValSerMetValAlaThrValIleGInSerLeuProIle 280
DB 815 ACCTTGCAAGAGTACCAATGGTCTCAATGGTCCGCAAGGTACTGAGTCCGCAACCA 874
QY 281 ThrAspGInLeuAspAsnIleSerValProLeuAlaThrIleLeuAlaAlaTyrLeu 300
DB 875 ACCGATCATTTAGACGACAAATATTTGCTTCTGCTGCTACATATTTTACGCTTATTA 934
QY 301 SerPheGlyTyr 304
DB 935 AGTTGCGATAT 946
RESULT 2
US-10-634-548-1
; Sequence 1, Application US/10634548
; Publication No. US20040045051A1
; GENERAL INFORMATION:

APPLICANT: No. US20040045051A1ris, Susan R
APPLICANT: Lincoln, Kim
APPLICANT: Abad, Mark Scott
APPLICANT: Eilers, Robert
APPLICANT: Hartsuiker, Karen Kindie
APPLICANT: Hirschberg, Joseph
APPLICANT: Karunanandaa, Balasubramanian
APPLICANT: Moshiri, Farhad
APPLICANT: Stein, Joshua C.
APPLICANT: Valentin, Henry B.
APPLICANT: Venkatesh, Tyamgondlu V.
TITLE OF INVENTION: Tocopherol biosynthesis related genes and used thereof
FILE REFERENCE: Ren-01-125
CURRENT APPLICATION NUMBER: US/10/634,548
CURRENT FILING DATE: 2003-08-05
PRIOR APPLICATION NUMBER: us 60/400,689
PRIOR FILING DATE: 2002-08-05
NUMBER OF SEQ ID NOS: 79
SOFTWARE: PatentIn version 3.2
SEQ ID NO 1
LENGTH: 1091
TYPE: DNA
ORGANISM: Arabidopsis thaliana
US-10-634-548-1
Alignment Scores:
Pred. No.: 1,86e-177 Length: 1091
Score: 1536.00 Matches: 304
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 7
US-10-634-548-2 (1-304) x US-10-634-548-1 (1-1091)
QY 1 MetAlaAlaThrLeuProLeuSerProIleAenHISGInLeuCyAArgPheGlyValAsn 20
DB 108 ATGGACGAACTTACTCTATCTCCGATCATCATAGTGTGTGCGGTCCGGAAAC 167
QY 21 SerLeuThrThrHISArgPheCysSerProGlyPheLeuIleSerSerProCysPheIle 40
DB 168 TCTTGGACGACTCACCGGCTTCTGTCTCTGCTTCTGATTCCTCTCTGTTCA 227
QY 41 GlyLeuThrGlyMetGlySerAlaThrGInLeuArgAlaArgArgSerLeuIleSerSer 60
DB 228 GGTGGACCGAAATGGGCTCTGCTACTAGTACGTGCTGCTGCTTCTGATCTTCTTCA 287
QY 61 AlaValAlaThrAsnSerLeuLeuHISAspValGlyAlaThrValAlaValLeuGlyGly 80
DB 288 GCAGTTGCGAAGAAATTCGCTGTGATGACGTCGAGCCACCGTGGCAGTGTGTGA 347
QY 81 AlaTyrAlaLeuValLeuSerPheGlyLeuThrIleValArgAsnValIleGInGInSer 100
DB 348 GCATACGGCGCTGTCTTAAGCTTCAGAGCTTCACCAAGCGAAAGTCATTCACAGAGT 407
QY 101 LeuSerArgGlyLeuValHISIleLeuSerGlyLeuLeuPheValLeuAlaTrpProIle 120
DB 408 TTGACGAGAAAGCTGTGATATCTCTCAGGCTGCTTCTGTTGCTGCTGCGCAATC 467
QY 121 PheSerGlySerThrGInAlaArgTyrPheAlaAlaPheValProLeuValLeuGlyLeu 140
DB 468 TTCACGCGATCGACCGAGGCTCGATATCTTGGCTTGTTCCTGTAGTGAATGGCTTA 527
QY 141 ArgLeuValIleAsnGlyLeuSerIleSerProAsnSerMetLeuIleYSerValThr 160
DB 528 AGCGTGTATTAACGAGCTATCCATTCCTCCAAATTCATCTAATCAATCCGTACA 587
QY 161 ArgGInGlyArgAlaGInGInLeuLeuGlyGlyProLeuPheTyrValLeuAlaLeuLeu 180
DB 588 AAGAGAGGAGAGAGAGAGAGTTCCTTAAGGCTTCTTCTTCTTCAAGTCTTCTT 647
QY 181 PheSerAlaValPhePheTrpArgIleUserProIleGlyMetCIIeSerLeuAlaMetMet 200

Db 648 TCTCTGCGGTTTCTTCTGAGAGAGTCTCTATCGATGATCTCGTTAGCAATGANG 707
Qy 201 CysGlyGlyAspGlyTLeuAlaSerPheGlyValGlyPheGlySerThrIlePro 220
Db 708 TGTGTGGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 767
Qy 221 TyrAspProArgLysSerTrpAlaGlySerIleSerMetPheIlePheGlyPhePhe 240
Db 768 TACAAACCAAGAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 827
Qy 241 SerIleAlaLeuLeuTyrTyrTyrSerSerLeuGlyTyrLeuHisMetAsnTrpIle 260
Db 828 TCCATCGCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 887
Qy 261 ThrLeuGlnArgValAlaMetValSerMetValAlaThrValAlaGluSerLeuPro 280
Db 888 ACCCTGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 947
Qy 281 ThrAspGlnLeuAspAspAsnIleSerValProLeuAlaThrIleLeuAlaIleTyr 300
Db 948 ACCGATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1007
Qy 301 SerPheGlyTyr 304
Db 1008 AGTTTCGGATAT 1019

RESULT 3

US-10-425-114-14606
; Sequence 14606, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jindong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses thereof for Plant Improvement
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 14606
; LENGTH: 980
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB23-048-A2_FLI
US-10-425-114-14606

Alignment Scores:

Pred. No.: 6,166-176 Length: 980
Score: 1523.00 Matches: 301
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 59.2% Indels: 0
DB: 7 Gaps: 0

US-10-634-548-2 (1-304) x US-10-425-114-14606 (1-980)

Qy 4 ThrLeuProLeuSerProIleAsnIleGlnLeuGlyAspPheGlyValAsnSerLeuThr 23
Db 3 ACCCTACCTCTATCTCCATCATCATCATCATCATCATCATCATCATCATCATCATCAT 62
Qy 24 ThrIleAspPheCysSerProGlyPheLeuIleSerSerProCysPheIleGlyLeuThr 43
Db 63 ACTACCGGTTCTGTCTCTCGGCTTCTGATTTCTTCTCTCTCTCTCTCTCTCTCTCT 122
Qy 44 GlyMetGlySerAlaThrGlnLeuArgAlaArgArgSerLeuIleSerSerAlaValAla 63
Db 123 GGAATGGGCTCTGCTACTCATGATGATGATGATGATGATGATGATGATGATGATG 182

Qy 64 ThrAsnSerLeuLeuHisAspValAlaThrValAlaValLeuGlyValIleTyrAla 83
Db 183 ACCAATTCGCTGTTCTCAAGACGTCCGAGCCACCGTGGAGTCTGTTGGAGCATACGG 242
Qy 84 LeuValIleSerPheGlySerLeuThrIleAspAsnValIleGlnIleSerLeuSerArg 103
Db 243 CTGTCTTAAGCTTGAAGGCTCTCAACGAGGAAACGTCATTCAACAGAGTTGAGCAGA 302
Qy 104 LysLeuValHisIleLeuSerGlyLeuLeuPheValIleuAlaTrpProIlePheSerGly 123
Db 303 AAGCTTGGCAATATCTCTCAAGTCTGCTTTCGTAAGTCTGCTGCTGCTGCTGCTGCT 362
Qy 124 SerThrGlnAlaArgTyrPheAlaIlePheValProLeuValAsnGlyLeuArgLeuVal 143
Db 363 TCGACCGAGGCTCGATCTTCTGCTGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 422
Qy 144 IleAsnGlyLeuSerIleSerProAsnSerMetLeuIleLysSerValThrArgGly 163
Db 423 ATTAACGAGCTATCATTTCCCAATTCGATGATCAATCCGTCACAGAGAGGG 482
Qy 164 ArgAlaGluGluLeuLeuGlyProLeuPheTyrValLeuAlaLeuLeuPheSerAla 183
Db 483 AGACGAGAGAGTCTTAAAGTCTTGTGCTTAAAGTCTTAAAGTCTTAAAGTCTTAAAG 542
Qy 184 ValPhePheTyrArgGluSerProIleGlyMetIleSerLeuAlaMetMetCysGlyGly 203
Db 543 GTTTCCTCTGAGAGAGTCTCTATCGATGATGATGATGATGATGATGATGATGATGAT 602
Qy 204 AspGlyIleAlaAspIleMetGlyArgLysPheGlySerThrIleIleProTyrAsnPro 223
Db 603 GATGGAATAGCTGATATATATGAGAGCTTAAGTTGGGTCACTAATATCTTACAAACCA 662
Qy 224 ArgLysSerTrpAlaGlySerIleSerMetPheIlePheGlyPhePheIleSerIleAla 243
Db 663 AGAAAGAGTGTGGCGAGAGATCTCCATGTTCACTTCTGCTTCTTCAATCTCCATCGCA 722
Qy 244 LeuLeuTyrTyrTyrSerSerLeuGlyTyrLeuHisMetAsnTrpGluThrIleGln 263
Db 723 TTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTT 782
Qy 264 ArgValAlaMetValSerMetValAlaThrValAlaGluSerLeuProIleThrAspGln 283
Db 783 AGAGTACGATAGTCTCATATGATGATGATGATGATGATGATGATGATGATGATGAT 842
Qy 284 LeuAspAspAsnIleSerValProLeuAlaThrIleLeuAlaIleTyrLeuSerPheGly 303
Db 843 TTAGACGACAAATATTTCTGCTCTGCTCTACTTATTTAGCTGCTTATTTAAGTTCCGA 902
Qy 304 Tyr 304
Db 903 TAT 905

RESULT 4

US-10-634-548-3
; Sequence 3, Application US/10634548
; Publication No. US20040045051A1
; GENERAL INFORMATION:
; APPLICANT: No. US20040045051A1, Susan R
; APPLICANT: Lincoln, Kim
; APPLICANT: Abad, Mark Scott
; APPLICANT: Eilers, Robert
; APPLICANT: Hartsuyker, Karen Kindle
; APPLICANT: Hirschberg, Joseph
; APPLICANT: Karunanandaa, Balasubramanian
; APPLICANT: Moshiri, Farhad
; APPLICANT: Stein, Joshua C.
; APPLICANT: Valentini, Henry B.
; APPLICANT: Venkatesh, Tyamagondlu V.
; TITLE OF INVENTION: Tocopherol Biosynthesis related genes and used thereof
; FILE REFERENCE: Ren-01-15
; CURRENT APPLICATION NUMBER: US/10/634,548
; CURRENT FILING DATE: 2003-08-05
; PRIOR APPLICATION NUMBER: us 60/400,689

PRIOR FILING DATE: 2002-08-05
 NUMBER OF SEQ ID NOS: 79
 SOFTWARE: PatentIn version 3.2
 SEQ ID NO 3
 LENGTH: 1897
 TYPE: DNA
 ORGANISM: Arabidopsis thaliana
 US-10-634-548-3

Alignment Scores:
 Pred. No.: 3,226-146 Length: 1897
 Score: 1284.50 Matches: 301
 Percent Similarity: 52.48 Conservative: 0
 Best Local Similarity: 52.48 Mismatches: 3
 Query Match: 83.64 Indels: 272
 DB: 7 Gaps: 5

US-10-634-548-2 (1-304) x US-10-634-548-3 (1-1897)

QY	1	MetAlaIaThrLeuProleuSerProIleAenHISGInLeuCyArgPheGlyAenAan	20
DB	108	ATGGACGACCACTTACCTTATCTCCGATCAATCATCACTGTCGCTTCGGGAACAC	167
QY	21	SerLeuThrThrHISArgPheCySerProGlyPheLeuIleSerSerProCyArgPheIle	40
DB	168	TCTTGAAGACTCACCGGCTTCTCTCCGCTTCCTGATTTCTTCTCTGTTTCATT	227
QY	41	GlyLeuThrGlyMetGlySerAlaThrGlnLeuArgAlaArgArgSerLeuIleSerSer	60
DB	228	GCTTGGACCGAATGGGCTCTGCTCACTCACTTACGTCGTCGTCCTGATCTCTTCA	287
QY	61	AlaValAlaThrAsnSerLeuLeuHISAspValGlyAlaThrValAlaValLeuGlyGly	80
DB	288	GCAATGGGACCAATTCCTGCTGTCATGACGTCGAGCCACCGTGGCAGCTTGCTGGA	347
QY	81	AlaTyAlaLeuValLeuSerPheGlySerLeuThrTyAspAsnValIleGln-----	98
DB	348	GCATACGGCTTGTCTTAAGCTTCGAGAGTCCACCAAGCAAAAGTCATTCACAAGGTC	407
QY	98	-----	98
DB	408	TCTTAATAATCGTTTAACTTATTCACACAATTTCTCCGTTTACAATTCAGTTTATTCG	467
QY	98	-----	98
DB	468	AACACCACTATATGTTGAAAGAGTTCTCAAGTTGTGTTGCAGTACTCATTAGAA	527
QY	98	-----	98
DB	528	ACAATGATTAAGCCTAGAAATTTGTGTAATTAAGTTTTCATTCTGAAATTTTATTA	587
QY	98	-----	98
DB	588	GAAATGTAACACTTAGTAAGCAGTATACACTTATCATGACCAATCGTAAGCGGA	647
QY	98	-----	98
DB	648	CAAGAACAAGTGTCCAAAATTTTACCGCTTATATATGTAACACTTTCTCAACTC	707
QY	98	-----	98
DB	708	CCTTTAACTATCCGTAATCGCTACCGCTAAACATATACCGTTCCTTGTGTTAACA	767
QY	99	-----GlnSerLeuSerAr	103
DB	768	AAGTAAGAAAGAGAAACAATTAATTGATGTTTATATGATGAGCAAGTTTGAAGAG	827
QY	103	GlySerLeuValHISLeuSerGlyLeuLeuPheValLeuAlaTrpProIle-----	120
DB	828	AAAGCTTGTCATATCTCTCAGGTCTGCTTTTCGACTTGGCTGSCCAATCTTCAGGA	887
QY	120	-----	120

DB	888	TTGCTTCTCTATATGTTGTAATCTCTCTGATCACTTTTAAACATGATAGACTTCTGA	947
QY	121	-----PheSerGlySerThrGlnAlaArgTyPheAlaAlaPh	133
DB	948	TTTCTTTTACTATCTTTTAAGTTTAAAGGATTCACAGGCTGATCTTGTGCTCTT	1007
QY	133	ValAlaProLeuValAlaGlyLeuArgValIleAlaGlyLeuSerIleSerProAenSe	153
DB	1008	TGTTCCGTTAGTAATGGCTTAAAGCTTGTTATTAACGACTATCAATTTCCCAAAATTC	1067
QY	153	MetLeuIleIleSerValThrArgGlyGlyArgAlaGlu-----	166
DB	1068	GATGCTATCAAAATCCGTCACAAAGAGAGAGAGAGTAAGTTGTCTAGTTTCTTTT	1127
QY	166	-----	166
DB	1128	TCCAACTTGATATGATTTTCAACAATCTGATTAACACTTCTGTTTCCAAACATCA	1187
QY	167	-----GlnLeuLeuIleGlyProleuPheTyValLeuAlaLeuLeuPheSerAlaValPh	185
DB	1188	CAGAGAGTTGCTTAAAGTCTTGTCTACGTTCTAGCTCTCTTCTCTGCGGCTTTT	1247
QY	185	ePheTyArgGlySerProIleGlyMetIleSerLeuAlaMetMetCyGlyGlyAspGly	205
DB	1248	CTTCTGAGAGAGTCTCTATCGGTATGATCTCTGACAAATGATGTGTGGGATGG	1307
QY	205	YIle-----	206
DB	1308	TAA-ATTTCTGTCAAGTACTCTATTAATTAATTAATTAATTAATTAATTAATTAATTA	1366
QY	207	-----AlaAspIleMetGlyAr	212
DB	1367	ATGACTAATAAGTGTGATCAATATGCTATGATAGAAATAGCTGATTAATAGGAGCG	1426
QY	212	GlyPheGlySerThrTyIleProTyAspProArgIleSerTyPheIleGlySerIleSe	232
DB	1427	TAAAGTTGGTCACTAAGATTAACCTTACCAACCAAGAAAGATTAAGGAGAGCATCTC	1486
QY	232	MetPheIlePheGlyPhePheIleSerIleAla-----	243
DB	1487	CATGTCATCTTCGGCTTCTTCACTCATCTCATGCG-CTAAATATTAACAATCCCATATT	1545
QY	243	-----	243
DB	1546	AATCATCAAAATGTCTCTCTCTGACGAGAAAGTCTTAAGAATGAGATGAGTTGC	1605
QY	244	-----LeuLeuTyTyTyTySerSerIle	251
DB	1606	TACTAAACCTTAACGCTTTCTTTGTAAATTTTGCAGATTAATCTTATTAATCTCAAGCT	1665
QY	251	GlyTyLeuHISMetAenTyPheTyThrThrLeuGlnArgValAlaMetValSerMetVa	271
DB	1666	TGGTACTTTCATGAATGAGCTGGAAAGACCTTGACAGAGATGACAAATGCTCAATGCT	1725
QY	271	AlaThrValValGlySerLeuProIleThrAspGlnLeuAspAsnIleSerValPr	291
DB	1726	CGCCACGGTAGTGCAGTGCCTAACCATCAACGATCAATTAAGACGAATATTCGGTTC	1785
QY	291	AlaThrValLeuAlaIleTyLeuSerPheGlyTyAr	304
DB	1786	TCTGCTACTATTTAAGCTGTATTAAGTTTGGGATAT	1825

RESULT 5
 US-10-425-114-29129
 ; Sequence 29129, Application US/10425114
 ; Publication No. US20040034688A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Liu, Jindong
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Kovalic, David K.
 ; APPLICANT: Screen, Steven E
 ; APPLICANT: Tabaska, Jack B
 ; APPLICANT: Cao, Yongwei

```

; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 29129
; LENGTH: 1062
; TYPE: DNA
; ORGANISM: Brassica napus
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB82-020-E9_FLI
US-10-425-114-29129

Alignment Scores:
Pred. No.: 2,366-144 Length: 1062
Score: 1266.00 Matches: 250
Percent Similarity: 89.3% Conservative: 24
Best Local Similarity: 81.4% Mismatches: 27
Query Match: 82.4% Indels: 6
Gaps: 2

US-10-634-548-2 (1-304) x US-10-425-114-29129 (1-1062)

QY 1 MetAlaAlaThrLeuProLeuSerProIleAenH1sGlnLeuCyArgPheGlyAenAen 20
Db 46 ATGGCGGAGCGCTTATCTTATCTCCGTTACCATCACTATATGTCGATAGCAACAG 105
QY 21 -----SerLeuThrThriAargPheCySerProGlyPheLeuIleSerSerPro 37
Db 106 TTCTGATTAACGCTATGACTCCCGGATCTGTGCGCA-----GTTCTTCGCCC 156
QY 38 CysPheIleGlyLeuThGlyMetGlySerAlaThrGlnLeuAargAlaArgArgSerLeu 57
Db 157 TGTATCATCGCGGAGAAAGAAATCGAGTCTGCACTTACGAGTTCGCGCTCGCACCTCTG 216
QY 58 IleSerSerAlaValAlaThrAenSerLeuLeuH1sAepValGlyAlaThrValAlaVal 77
Db 217 ATCTCTTACGCGCTTCGATCATATATCTATGCTAGCTGAGAGCCAGCGTGCAGTT 276
QY 78 LeuGlyGlyAlaThrAlaLeuValIleAenGlyLeuSerPheGlyLeuThrIleAenValIle 97
Db 277 CTATAGTGGAGCTTACGCGCTTGTCTTACCTTCGAGAGTCTCAGAGAGGAGAGCTGAT 336
QY 98 GlnGlnSerLeuSerAargGlyLeuValH1sIleLeuSerGlyLeuLeuPheValLeuAla 117
Db 337 CCACAGAGATTGAGCAAGAAAGCTTGTGCAATATCTCAGGCTCTCTTTCGCGCTCTCG 396
QY 118 TrpProIlePheSerGlySerThGlnAlaArgTyrrPheAlaAlaPheValProLeuVal 137
Db 397 TGGCCATCTTCAGCGCATCAACGAGGCTCGATATCTTGTGCTTGTGTCTTTAGTG 456
QY 138 AenGlyLeuAargLeuValIleAenGlyLeuSerIleSerProAenSerMetLeuIleAys 157
Db 457 AATGGCCTTAAGGCTTGTGTCAACGCTTGTGCGCTCCCTCAACTCCACGCTAATCCA 516
QY 158 SerValThrAargGlyAargAlaGlnGlnLeuLeuValGlyProLeuPheTyrrValLeu 177
Db 517 TCCGTAATCAGGAGGAGGAGCAAGAGAGTGTCTTAAGGCTCAATGTTCTTACCTTCTA 576
QY 178 AlaLeuLeuPheSerAlaValPhePheTrpAargGlySerProIleGlyMetIleSerLeu 197
Db 577 GCCCTTCTAGTTCGCGAGTTTCTTCTGAGAGATTTCTCTCAACCGGATATGATATCGCTG 636
QY 198 AlaMetMetCyArgGlyAargGlyIleAlaAepIleMetGlyAargGlyPheGlySerThr 217
Db 637 GCAATGATGTGTGGCGAGCGAGATAGCTGATATCATGCGAGCGTAAGACGATCATAC 696
QY 218 LysIleProTyrrAenProAargGlySerTrpAlaGlySerIleSerMetPheIlePheGly 237
Db 697 AAGATTAACCTTACCAACCAAGAAAGAGCTTGGCGGAGAGCATCTCATGTTCAATTTCGCG 756
QY 238 PhePheIleSerIleAlaLeuLeuTyrrTyrrSerSerLeuGlyTyrrIleuH1sMetAen 257

```

```

Db 757 TTCTTCACTTCATCGGATTACTTACTTACTTACTAAGCGCTAGGCTATCTTCACTGAC 816
QY 258 TrpGlyThrThrLeuGlnAargValAlaMetValSerMetValAlaThrValAlaGlySer 277
Db 817 TGGGAAACACACTTATACAGAGTCCGATTTGCTCATTTGCTGCTACGCTGTGGAGTCA 876
QY 278 LeuProIleThrAargGlnLeuAenAenIleSerValProLeuAlaThrIleLeuAla 297
Db 877 CTACCATCAACCGATCAATATACGACAAACGTTTGGTTCTCTGCTACTATTTTGCT 936
QY 298 AlaTyrrLeuSerPheGlyTyrr 304
Db 937 GCTTACCTTAAGTTTGGATAT 957

RESULT 6
US-10-425-114-20410
; Sequence 20410, Application US/10425114
; Publication No. US2004003488A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jindong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven B
; APPLICANT: Tabaska, Jack B
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 20410
; LENGTH: 1007
; TYPE: DNA
; ORGANISM: Gossypium hirsutum
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3196-048-All_FLI
US-10-425-114-20410

Alignment Scores:
Pred. No.: 8,196-96 Length: 1007
Score: 870.00 Matches: 174
Percent Similarity: 73.3% Conservative: 51
Best Local Similarity: 56.7% Mismatches: 72
Query Match: 56.6% Indels: 10
Gaps: 2

US-10-634-548-2 (1-304) x US-10-425-114-20410 (1-1007)

QY 2 AlaAlaThrLeuProLeuSerProIleAenH1sGlnLeuCyArgPheGlyAenAenSer 21
Db 18 GCTGTACACAGAGGCTCTTC-----TTATCTTCACTCATCAATC 59
QY 22 LeuThrThriAargPheCySer-----ProGlyPheLeuIleSerSerPro 37
Db 60 TTAAAGCCGCAAGCTATATCCGCGATTTCTCTCTCCCGCTTCTTCTCTCTCTCT 119
QY 38 CysPheIleGlyLeuThGlyMetGlySerAlaThrGlnLeuAargAlaAargSerLeu 57
Db 120 CTCATCCCAACAGGCTCCGTTTCCCTATTTCTTACGCGGCGCCCAACGCGCACCGCA 179
QY 58 IleSerSerAlaValAlaThrAenSerLeuLeuH1sAepValGlyAlaThrValAlaVal 77
Db 180 CTCACGCAACCGCTGTAAACGCTCTTATCTTCGAGATACCGCGCTTCTGCGCTGTGC 239
QY 78 LeuGlyGlyAlaThrAlaLeuValIleAenSerPheGlyLeuThrIleAenValIle 97
Db 240 TTGTGCTGCGCTTATGCTCTGCTCTTCACTTCACTTCTCATCTCAAAAGAGCTATT 299
QY 98 GlnGlnSerLeuSerAargGlyLeuValH1sIleLeuSerGlyLeuLeuPheValLeuAla 117
Db 300 CAGCAATTTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAG 359

```

Alignment Scores:	
pred. No.:	1 94e-86
Score:	799.00
Percent Similarity:	70.9%
Best Local Similarity:	51.8%
Query Match:	51.7%
DB:	7
	Gaps: 3
	length: 1085
	Matches: 162
	Conservative: 60
	Mismatches: 78
	Indels: 13

RESULT 8
US-10-437-963-77300
; Sequence 77300, Application US/10437962
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.


```

Db      897 GCCTATCTGTATTGGCTAC 917
|||||
RESULT 11
US-10-425-115-118753
/ Sequence 118753, Application US/10425115
/ Publication No. US20040214272A1
/ GENERAL INFORMATION:
/ APPLICANT: La Rosa, Thomas J.
/ APPLICANT: Kovalic, David K.
/ APPLICANT: Zhou, Yihua
/ APPLICANT: Cao, Yongwei
/ TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
/ FILE OF INVENTION: Plants
/ FILE REFERENCE: 38-21(53222)B
/ CURRENT APPLICATION NUMBER: US/10/425,115
/ CURRENT FILING DATE: 2003-04-28
/ NUMBER OF SEQ ID NOS: 369326
/ SEQ ID NO 118753
/ LENGTH: 1123
/ TYPE: DNA
/ ORGANISM: Zea mays
/ FEATURE:
/ OTHER INFORMATION: Clone ID: MRY577_39790C.1
US-10-425-115-118753

Alignment Scores:
Pred. No.:      2.5e-84      Length:      1123
Score:          777.00      Matches:     152
Percent Similarity: 73.0%      Conservative: 56
Best Local Similarity: 53.3%      Mismatches: 65
Query Match:    50.6%      Indels:     12
DB:             8          Gaps:          3

US-10-634-548-2 (1-304) x US-10-425-115-118753 (1-1123)
Qy      29 SerProGlyPheLeuile-----SerSerProGlyPheLeuThrglyMetgly 46
      |||||
Db      129 TCCCCCACTGCTCTGCTGCTGCGGCTGCTCCCGCCAGCCGCGCGCTGCGGCTTGC 188
      |||||
Qy      47 SerAlaThrGlnLeuArgAlaArgSerLeuileSer----- 59
      |||||
Db      189 CCGGGAGCTCATCGCG--AGCGGCTCTGCTCGGCGTGGGACCCCGCGTGGCG 245
      |||||
Qy      60 -----SerAlaValAlaThrAsnSerLeuHisaPvalGlyAlaThrAlaVal 77
      |||||
Db      246 GCGCTGGCGCGCGCGCGCGCGCGCGCTGTCAGACGGGCGGTCACTGTCTCATC 305
      |||||
Qy      78 LeuGlyAlaTyrAlaLeuValLeuSerPheGluSerLeuThrLysArgAsnValile 97
      |||||
Db      306 ACCGCGCGGCTTACTCTCTTGGCGCGCTTTCGACGAGCTACCGAGCGCGGCTCATC 365
      |||||
Qy      98 GlnGlnSerLeuSerArgLysLeuValHisIleLeuSerGlyLeuLeuPheValLeuAla 117
      |||||
Db      366 GAAAGACCTTGACGAGGAGGTGTGACGCTGATCCGGGCTCTTTCATGTCATCT 425
      |||||
Qy      118 TrpProIlePheSerGlySerThrGlnAlaArgLysPheAlaAlaPheValProLeuVal 137
      |||||
Db      426 TGGCCCTGTCAGCAATTCGACAGAGACCGGTATTCGCGCGGTGTCCCGTCTCG 485
      |||||
Qy      138 AsnGlyLeuArgLeuValIleAsnGlyLeuSerIleSerProAsnSerMetLeuIleLys 157
      |||||
Db      486 AACTCATGAGGCTCTGATATATGACTCCGCTCTACACTGATGAAGCTGTGATANA 545
      |||||
Qy      158 SerValThrArgGlnGlyArgAlaGlnLeuLeuLysGlyProLeuPheTyrValLeu 177
      |||||
Db      546 TCAGTACACGCTGAGGAGAAACACAGAGATGTCGAGAGGTCACCTCATATATCTCG 605
      |||||
Qy      178 AlaLeuLeuPheSerAlaValPhePheTyrPargLysSerProIleGlyMetIleSerLeu 197
      |||||
Db      606 GTGCTGCTGTCAGGCTTATGCTTCTGCGCTGAGTCCCGCATCGGATGCTCTCC 665
      |||||
Qy      198 AlaMetMetCysGlyGlyAspGlyIleAlaAspIleMetGlyArgLysPheGlySerThr 217
      |||||

```

```

Db      666 TCGATGATGAGCGGTGGCGATGTTTCTGACATGTTGGAGAGGATATGGCTCAAG 725
      |||||
Qy      218 LysIleProTyrAsnProArgLysSerTrpAlaGlySerIleSerMetPheIlePheGly 237
      |||||
Db      726 AAGCTGCATTCATTCGAGAGAGAGCTGGCGCGGAGACATCTCGATTCATTTCTG 785
      |||||
Qy      238 PhePheIleSerIleAlaLeuLeuTyrTyrIleSerLeuGlyTyrLeuHisaMetAsn 257
      |||||
Db      786 TTCCTGCTGCGCATATGCTTACTTCTTCAAGCTGGGATTCATGATGATCTTATC 845
      |||||
Qy      258 TrpGlnThrThrLeuGlnArgValAlaMetValSerMetValAlaThrValGlnSer 277
      |||||
Db      846 TCGGAGAGGACACTTGTGAAGCTGCGCTTGTGACATAGACAGACAGTATGAGAGTC 905
      |||||
Qy      278 LeuProIleThrAspGlnLeuAspAsnIleSerValProLeuAlaThrIleLeuAla 297
      |||||
Db      906 GTTCTGTGACCGAAGTTGATGATGACAATATCTGCTTGGCCACCATGCTGATA 965
      |||||
Qy      298 AlaTyrLeuSerPhe 302
      |||||
Db      966 GCTTTCTCTTGT 980
      |||||

RESULT 12
US-10-425-114-23255
/ Sequence 23255, Application US/10425114
/ Publication No. US20040034888A1
/ GENERAL INFORMATION:
/ APPLICANT: Liu, Jingdong
/ APPLICANT: Zhou, Yihua
/ APPLICANT: Kovalic, David K.
/ APPLICANT: Screen, Steven E
/ APPLICANT: Tabaska, Jack E
/ APPLICANT: Cao, Yongwei
/ TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
/ FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
/ FILE REFERENCE: 38-21(5313)B
/ CURRENT APPLICATION NUMBER: US/10/425,114
/ CURRENT FILING DATE: 2003-04-28
/ NUMBER OF SEQ ID NOS: 73128
/ SEQ ID NO 23255
/ LENGTH: 798
/ TYPE: DNA
/ ORGANISM: Zea mays
/ FEATURE:
/ OTHER INFORMATION: Clone ID: LIB3595-006-H9_FLI
US-10-425-114-23255

Alignment Scores:
Pred. No.:      1.91e-84      Length:      798
Score:          776.00      Matches:     143
Percent Similarity: 80.2%      Conservative: 52
Best Local Similarity: 58.8%      Mismatches: 48
Query Match:    50.5%      Indels:     0
DB:             7          Gaps:          0

US-10-634-548-2 (1-304) x US-10-425-114-23255 (1-798)
Qy      60 SerAlaValAlaThrAsnSerLeuLeuHisaPvalGlyAlaThrValAlaValLeuGly 79
      |||||
Db      1 GCGGCGCGGCGCGCGCGGCTGTCAGACCGAGCGGTCACTGCTCATACCGCC 60
      |||||
Qy      80 GlyAlaTyrAlaLeuValLeuSerPheGluSerLeuThrLysArgAsnValIleGln 99
      |||||
Db      61 GCGGCTTACTCTTGTGCGGCTTTCGACAGCTCACCGGCGCGGCTCATCGAAG 120
      |||||
Qy      100 SerLeuSerArgLysLeuValHisIleLeuSerGlyLeuLeuPheValLeuAlaTrpPro 119
      |||||
Db      121 AGCTTACAGAGAGGTGTGACGCTGATCCGCGCTCTTTCATATCTTGGGCC 180
      |||||
Qy      120 IlePheSerGlySerThrGlnAlaArgLysPheAlaAlaPheValProLeuValAsnGly 139
      |||||
Db      181 CTGTTACGACATTCGACAGAGACCGGATTTCCCGGCTTCCCGTCTGAACTCC 240
      |||||

```

QY 140 LeuArgLeuValIleAsnGlyLeuSerIleSerProAsnSerMetLeuIleYSerVal 159
 Db 241 ATGAGAGCTTGAATATGAGACTCCGCTCTCACTGATGAGAGCTGTGTAATCAATG 300
 QY 160 ThrArgGluGlyArgAlaGluGluLeuLeuYsglyProLeuPheTyValIleuAlaLeu 179
 Db 301 ACAGGTGAAGGAAACAGAGGAATGCTGAGAGGTCCACTATTAATGCTCGGTGCTG 360
 QY 180 LeuPheSerAlaValPhePheTTPArgGluSerProIleGlyMetIleSerLeuAlaMet 199
 Db 361 CTGTTCAGGGTTTATGCTCTTCGGCGGTGAGTCCCATCGGAGATGCTCTTCGTCAGT 420
 QY 200 MetCysGlyGlyAspGlyIleAlaAspIleMetGlyArgLeuPheGlySerThrIleYsle 219
 Db 421 ATGACCGGTGCGAGTGGTTTCTCTCACTATTTGGAGAGAGATAGGCTCAGCGAAGCTG 480
 QY 220 ProTyAsnProArgIleYSerTTPAlaGlySerIleSerMetPheIlePheGlyPhePhe 239
 Db 481 CCATTCAATCGGAGAGAGAGCTGGGCGGAGCATCTCGATGTTCAATTTCTGGTTCTG 540
 QY 240 IleSerIleAlaLeuLeuTyTyTyTySerSerIleuGlyTyTyLeuHiMetAsnTTPGlu 259
 Db 541 CTGTCCGCGATGATGATGCTCTCACTCTCAAGCTGGGTTACATTCATTCATTCGGA 600
 QY 260 ThrThrLeuGlnArgValAlaMetValIleSerMetValAlaThrValAlaGluSerLeuPro 279
 Db 601 GAGGACCTTGTGAAGCTGCGCTGTGTGCACTAGCAGCAGCATGTGAGTCCCTTCT 660
 QY 280 IleThrAspGluLeuAspAspAsnIleSerValProLeuAlaThrIleLeuAlaIleTy 299
 Db 661 GTGACCGAAGTTGTGATGACAACATATCTGTTCTTGGCCACCATGCTGTAGCTTTT 720
 QY 300 LeuSerPhe 302
 Db 721 CTCTGTTT 729
 RESULT 13
 US-10-425-114-26551
 / Sequence 26551, Application US/10425114
 / Publication No. US20040034888A1
 / GENERAL INFORMATION:
 / APPLICANT: Liu, Jindong
 / APPLICANT: Zhou, Yihua
 / APPLICANT: Kovalic, David K.
 / APPLICANT: Screen, Steven E
 / APPLICANT: Tabaska, Jack E
 / APPLICANT: Cao, Yongwei
 / TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
 / FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
 / FILE REFERENCE: 38-21(5313)B
 / CURRENT APPLICATION NUMBER: US/10/425, 114
 / CURRENT FILING DATE: 2003-04-28
 / NUMBER OF SEQ ID NOS: 73128
 / SEQ ID NO 26551
 / LENGTH: 976
 / TYPE: DNA
 / ORGANISM: Triticum aestivum
 / FEATURE:
 / OTHER INFORMATION: Clone ID: LIB4312-010-Cl1_FLI
 / US-10-425-114-26551
 Alignment Scores:
 Pred. No.: 6,17e-84 Length: 976
 Score: 773.00 Matches: 149
 Percent Similarity: 70.9% Conservative: 56
 Best Local Similarity: 51.6% Mismatches: 66
 Query Match: 50.3% Indels: 18
 DB: 7 Gaps: 1
 US-10-634-548-2 (1-304) x US-10-425-114-26551 (1-976)
 QY 33 LeuIleSerSerProCysPheIleGlyLeuThrGlyMetCysAlaThrGlnLeuArg 52

Db 60 CTCCTCTCTCCCGACCTGCTGCTGCGGCTTCACTCCGCGCCGACCTGCGC 119
 QY 53 AlaArgArg----- 55
 Db 120 GCCCGAGAGCCGAGCCGCTGCTGCGCGCCGCGCCGCGCGCGCGCGCGCG 179
 QY 56 ---SerLeuIleSerSerAlaValAlaThrAsnSerLeuLeuHiAspValGlyAlaThr 74
 Db 180 CTGGCGGCGCTGCGCTGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 239
 QY 75 ValAlaValLeuGlyAlaTyTyTyTyAlaLeuValLeuSerPheGluSerLeuThrIleArg 94
 Db 240 CTGCTCGTCACCGCGCGCGCGCTTCTCTCTGCGCGCGCGCGCGCGCGCG 299
 QY 95 AsnValIleGlnIleSerLeuSerArgGlySerValIleIleLeuSerGlyLeuLeuPhe 114
 Db 300 CGCCTCGTCACAGAGCTTGAAGAGAGAGTGTGATGTGCTATCCGGGGTCTTTTTC 359
 QY 115 ValLeuAlaTTPProIlePheSerGlySerThrGlnAlaArgTyTyPheAlaAlaPheVal 134
 Db 360 ATGGCTTCATGCGCACTCTTCAAGCAATTCACAGTGACCGGTTCTTCCGAGCGGTAGT 419
 QY 135 ProLeuValAsnGlyLeuArgLeuValIleAsnGlyLeuSerIleSerProAsnSerMet 154
 Db 420 CCGTTCCTGAACCTGCGTGAAGCTTCTCACTACGCGGCTCGCTTCTACTCCGACGAGACT 479
 QY 155 LeuIleYSerSerValIleArgGluGlyArgAlaGluGluLeuLeuLeuGlyProLeuPhe 174
 Db 480 CTAGTAATATCGTGAAGCCCGTGAAGAGAAACAGAGAAATTCCTGCGAGGCCCTCTGAG 539
 QY 175 TyTyValLeuAlaLeuLeuPheSerAlaValPhePheTTPArgGluSerProIleGlyMet 194
 Db 540 TATGTCATCGTCACTGATCATTTGTTAGTCTTTTGGCGGAGTCCCGCATCGGAGATC 599
 QY 195 IleSerLeuAlaMetMetCysGlyGlyAspGlyIleAlaAspIleMetGlyArgIlePhe 214
 Db 600 GTTTCCTTGTGATGATAGCGGTGTGACGCGCTTGTGACATTTGTGGAGAAAGTTC 659
 QY 215 GlySerThrTyIleProTyAsnProArgIleYSerTTPAlaGlySerIleSerMetPhe 234
 Db 660 GCTTCGTCAGAGCTGCCATTCAACAGAGAGAGCTGGTCGAGAGTCCCGCATGTTC 719
 QY 235 IlePheGlyPhePheIleSerIleAlaLeuLeuTyTyTyTySerSerLeuGlyTyTyLeu 254
 Db 720 ATCTCGGTTCTCGCTATCTGACCTGATGCTGCTTCTCTGCTGCTGCTGCTGCTGCT 779
 QY 255 HiMetAsnTTPGluThrThrLeuGlnArgValAlaMetValIleSerMetValAlaThrVal 274
 Db 780 CATGTCACCTGGATCAGCGCTTGTGTAACGTGTTCTGCTGCGCTGCGACGCACTGTG 839
 QY 275 ValGluSerLeuProIleThrAspGluLeuAspAspAsnIleSerValProLeuAlaThr 294
 Db 840 GTGAGAGTATTCCTGTAAGTATGATTTGTGATGACATATCTCTGTTCCCTTGGCGCAC 899
 QY 295 IleLeuAlaAlaTyTyTyTySerPheGly 303
 Db 900 ATGTTGTAGCTTCTCTGTTTGGC 926
 RESULT 14
 US-10-425-114-26571
 / Sequence 26571, Application US/10425114
 / Publication No. US20040034888A1
 / GENERAL INFORMATION:
 / APPLICANT: Liu, Jindong
 / APPLICANT: Zhou, Yihua
 / APPLICANT: Kovalic, David K.
 / APPLICANT: Screen, Steven E
 / APPLICANT: Tabaska, Jack E
 / APPLICANT: Cao, Yongwei
 / TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
 / FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
 / FILE REFERENCE: 38-21(5313)B

```

/ CURRENT APPLICATION NUMBER: US/10/425,114
/ CURRENT FILING DATE: 2003-04-28
/ NUMBER OF SEQ ID NOS: 73128
/ SEQ ID NO 25671
/ LENGTH: 723
/ TYPE: DNA
/ ORGANISM: Zea mays
/ FEATURE:
/ OTHER INFORMATION: Clone ID: LIB3957-022-E3_FLI
US-10-425-114-25671

Alignment Scores:
Pred. No.: 1,9e-81 Length: 723
Score: 751.00 Matches: 137
Percent Similarity: 83.0% Conservative: 48
Best Local Similarity: 61.4% Mismatches: 38
Query Match: 48.9% Indels: 0
Gaps: 0

US-10-634-548-2 (1-304) x US-10-425-114-25671 (1-723)

QY 80 GYAAATYRAALeUValLeuSerPheGluSerLeuThrLyArgAsnValIleGlnGln 99
Db 7 GGCGCTACTCTCTTGTGGCGCTTCGACGAGCTCACCGAGCGCGCTCATCGAAAG 66
QY 100 SerLeuSerArgLyLeuValHisIleLeuSerGlyLeuLeuPheValLeuAlaTrpPro 119
Db 67 AGCTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 126
QY 120 IlePheSerGlySerThrGluAlaArgTyPheAlaIlePheValProLeuValAsnGly 139
Db 127 CTGTTACAGAAATTCACAGAGAACCGATATTCGCGCGGTTCCTCCGTCGAACTCC 186
QY 140 LeuArgLeuValIleAsnGlyLeuSerIleSerProAsnSerMetLeuIleLysSerVal 159
Db 187 ATGAGCGCTCTGATATATGACTCCGCTCTACACTGATGAAGACTGTGTAATCAAGTG 246
QY 160 ThrArgGluGlyArgAlaGluGluLeuLeuLysGlyProLeuPheTyPheValLeuAlaLeu 179
Db 247 ACACGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 306
QY 180 LeuPheSerAlaValPhePheTrpArgGluSerProIleGlyMetIleSerLeuAlaMet 199
Db 307 CTGTTACAGAGTTTACGCTTCTTGGCGTGAAGTCCCATCGAGATGCTCTCTTGCATG 366
QY 200 MetCysGlyGlyAspGlyTyrIleAlaAspIleMetGlyArgLyPheGlySerThrIlySile 219
Db 367 ATGACCGGTGGCGAGTGTTCGCTGACATTTGTTGGAGAGAGATGAGTCAAGAGCTG 426
QY 220 ProTyraAnProArgLySerTrpAlaGlySerIleSerMetPheIlePheGlyPhePhe 239
Db 427 CCATTCAATCGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 486
QY 240 IleSerIleAlaLeuLeuTyTyrTyPheSerLeuGlyTyrLeuHisMetAsnTrpGlu 259
Db 487 CTGTCGGCATATATGCTTACTTCTCAAGCTCGGTTACATGATGTTATCTCGGAA 546
QY 260 ThrThrLeuGlnArgValAlaMetValSerMetValAlaThrValValGluSerLeuPro 279
Db 547 GAGGCACTTGTAACCTGGCGCTGTGTCATAGACGACAGAGTGAAGTGGCTTCT 606
QY 280 IleThrAspGlnLeuAspAspAsnIleSerValProLeuAlaThrIleLeuAlaIaTyP 299
Db 607 GTGACCGAAGTTGTGATGATCAACATCTGTTCTTGGCCACCATGCTGGTAGCTTTT 666
QY 300 LeuSerPhe 302
Db 667 CTCTTGTT 675

```

```

/ GENERAL INFORMATION:
/ APPLICANT: Liu, Jingdong
/ APPLICANT: Zhou, Yihua
/ APPLICANT: Kovalic, David K.
/ APPLICANT: Screen, Steven E
/ APPLICANT: Tabaska, Jack E
/ APPLICANT: Cao, Yongwei
/ TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
/ TITLE OF INVENTION: Plants and Uses thereof for Plant Improvement
/ FILE REFERENCE: 38-21 (5313)B
/ CURRENT APPLICATION NUMBER: US/10/425,114
/ CURRENT FILING DATE: 2003-04-28
/ NUMBER OF SEQ ID NOS: 73128
/ SEQ ID NO 26498
/ LENGTH: 753
/ TYPE: DNA
/ ORGANISM: Zea mays
/ FEATURE:
/ OTHER INFORMATION: Clone ID: LIB4185-055-C9_FLI
US-10-425-114-26498

Alignment Scores:
Pred. No.: 2,03e-81 Length: 753
Score: 751.00 Matches: 137
Percent Similarity: 83.0% Conservative: 48
Best Local Similarity: 61.4% Mismatches: 38
Query Match: 48.9% Indels: 0
Gaps: 0

US-10-634-548-2 (1-304) x US-10-425-114-26498 (1-753)

QY 80 GYAAATYRAALeUValLeuSerPheGluSerLeuThrLyArgAsnValIleGlnGln 99
Db 16 GGCGCTACTCTCTTGTGGCGCTTCGACGAGCTCACCGAGCGCGCTCATCGAAAG 75
QY 100 SerLeuSerArgLyLeuValHisIleLeuSerGlyLeuLeuPheValLeuAlaTrpPro 119
Db 76 AGCTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 135
QY 120 IlePheSerGlySerThrGluAlaArgTyPheAlaIlePheValProLeuValAsnGly 139
Db 136 CTGTTACAGAAATTCACAGAGAACCGATATTCGCGCGGTTCCTCCGTCGAACTCC 195
QY 140 LeuArgLeuValIleAsnGlyLeuSerIleSerProAsnSerMetLeuIleLysSerVal 159
Db 196 ATGAGCGCTCTGATATATGACTCCGCTCTACACTGATGAAGACTGTGTAATCAAGTG 255
QY 160 ThrArgGluGlyArgAlaGluGluLeuLeuLysGlyProLeuPheTyPheValLeuAlaLeu 179
Db 256 ACACGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 315
QY 180 LeuPheSerAlaValPhePheTrpArgGluSerProIleGlyMetIleSerLeuAlaMet 199
Db 316 CTGTTACAGAGTTTACGCTTCTTGGCGTGAAGTCCCATCGAGATGCTCTCTTGCATG 375
QY 200 MetCysGlyGlyAspGlyTyrIleAlaAspIleMetGlyArgLyPheGlySerThrIlySile 219
Db 376 ATGACCGGTGGCGAGTGTTCGCTGACATTTGTTGGAGAGAGATGAGTCAAGAGCTG 435
QY 220 ProTyraAnProArgLySerTrpAlaGlySerIleSerMetPheIlePheGlyPhePhe 239
Db 436 CCATTCAATCGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 495
QY 240 IleSerIleAlaLeuLeuTyTyrTyPheSerLeuGlyTyrLeuHisMetAsnTrpGlu 259
Db 496 CTGTCGGCATATATGCTTACTTCTCAAGCTCGGTTACATGATGTTATCTCGGAA 555
QY 260 ThrThrLeuGlnArgValAlaMetValSerMetValAlaThrValValGluSerLeuPro 279
Db 556 GAGGCACTTGTAACCTGGCGCTGTGTCATAGACGACAGAGTGAAGTGGCTTCT 615
QY 280 IleThrAspGlnLeuAspAspAsnIleSerValProLeuAlaThrIleLeuAlaIaTyP 299

```

Db 616 GTGACCGAAGTTGTAGATGACAATATCTGTTCCTTTGGCCACCATGCTGTAGCTTTT 675
Qy 300 LeuSerPhe 302
Db 676 CTTTGTGTTT 684

Search completed: March 16, 2006, 23:49:59
Job time : 915 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: March 16, 2006, 23:48:07 ; Search time 429 Seconds

(without alignments)
1652.304 Million cell updates/sec

Title: US-10-634-548-2

Perfect score: 1536

Sequence: 1 MAATLPSPINHQLCRFGNN.....DNISVPLATIAVLSFGY 304

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 8023312 seqs, 1165852854 residues

Total number of hits satisfying chosen parameters: 16046624

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:
-MODEL=frame+.p2n.model -DEV=xlh
-Q=/abs/ABSSWB.epool/US10634548/runat_16032006_084442_10504/app.query.fasta.1
-DB=Published.Applications.NA.New -QFMT=fastac -SUFPTX=p2n.rnpn -MINMATCH=0.1
-LOOPEL=0 -LOOPEXT=0 -UNITS=bits -SPART=1 -END=-1 -MATRX=bloms62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -HOST=abs802h
-USER=US10634548 @CGN 1 1 431 @runat_16032006_084442_10504 -NCPU=6 -ICPU=3
-NO MAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published.Applications.NA.New.*
1: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
2: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US05_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*
11: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*
12: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	777.5	50.6	1198	US-11-096-568A-19558	Sequence 19558, A
2	613	39.9	986	US-11-096-568A-1769	Sequence 1769, A
3	562.5	36.6	1048	US-11-096-568A-14811	Sequence 14811, A
4	475	30.9	481	US-11-096-568A-14317	Sequence 14317, A

5	152.5	9.9	978	US-10-932-182A-81375	Sequence 81375, A
6	152.5	9.9	978	US-10-932-182A-81375	Sequence 81375, A
7	150	9.8	525	US-10-932-182A-174337	Sequence 174337, A
8	150	9.8	525	US-10-932-182A-174337	Sequence 174337, A
9	143.5	9.3	804	US-10-932-182A-80837	Sequence 80837, A
10	143.5	9.3	804	US-10-932-182A-80837	Sequence 80837, A
11	143.5	9.3	1560	US-10-932-182A-1952	Sequence 1952, A
12	143.5	9.3	1560	US-10-932-182A-1952	Sequence 1952, A
13	115.5	7.5	7974	US-10-499-290-1	Sequence 1, App1
14	98.5	6.4	1152	US-10-793-626-2167	Sequence 2167, App
15	98.5	6.4	3108	US-10-793-626-3427	Sequence 3427, App
16	97	6.3	1296	US-11-092-140-5	Sequence 5, App1
17	97	6.3	2694	US-10-932-182A-78840	Sequence 78840, A
18	97	6.3	2694	US-10-932-182A-78840	Sequence 78840, A
19	95.5	6.2	867	US-11-194-246-495	Sequence 495, App
20	95.5	6.2	2037	US-11-096-568A-22288	Sequence 22288, A
21	95.5	6.2	2640	US-11-194-246-385	Sequence 385, App
22	95.5	6.2	2720	US-11-194-246-124	Sequence 124, App
23	95.5	6.2	8833	US-10-876-787-5	Sequence 5, App1
24	94.5	6.2	2085	US-10-485-517-391	Sequence 391, App
25	94.5	6.2	3840	US-10-793-626-4013	Sequence 4013, App
26	94.5	6.2	4235	US-10-793-626-4041	Sequence 4041, App
27	93.5	6.1	1230	US-10-793-626-2001	Sequence 2001, App
28	93.5	6.1	5121	US-10-793-626-2305	Sequence 2305, App
29	92.5	6.0	1784	US-11-052-554A-559	Sequence 559, App
30	91.5	5.9	962	US-11-096-568A-23660	Sequence 23660, A
31	90.5	5.9	1180	US-10-793-626-4349	Sequence 4349, App
32	90.5	5.9	1006	US-10-793-626-33210	Sequence 33210, A
33	90	5.9	1482	US-11-096-568A-33230	Sequence 33230, A
34	89	5.8	1482	US-10-467-657-4375	Sequence 4375, App
35	87.5	5.7	1345	US-09-925-065A-61537	Sequence 61537, A
36	87.5	5.7	1345	US-11-000-463-474	Sequence 474, App
37	87.5	5.7	1842	US-10-932-182A-76836	Sequence 76836, A
38	87.5	5.7	1842	US-10-932-182A-76836	Sequence 76836, A
39	87.5	5.7	1992	US-11-000-463-2	Sequence 2, App1
40	87.5	5.7	2928	US-11-103-957-48	Sequence 48, App1
41	87	5.7	1539	US-11-087-099-23	Sequence 23, App1
42	87	5.7	2076	US-11-098-686-4901	Sequence 9401, App
43	87	5.7	3117	US-10-793-626-4005	Sequence 4005, App
44	87	5.7	1457619	US-11-098-686-8739	Sequence 8739, App
45	86.5	5.6	663	US-10-821-234-648	Sequence 648, App

ALIGNMENTS

RESULT 1
US-11-096-568A-19558
; Sequence 19558, Application US/11096568A
; Publication No. US20060048240A1
GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; TITLE OF INVENTION: Theyxy
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
NUMBER OF SEQ ID NOS: 34471
SEQ ID NO 19558
; LENGTH: 1198
; TYPE: DNA
; ORGANISM: Zea mays subsp. mays
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)..(1198)
; OTHER INFORMATION: Ceres Seq. ID no. 12373616
US-11-096-568A-19558

Alignment Scores:
Pred. No.: 2,41e-72
Score: 777.50
Percent Similarity: 68.9%
Best Local Similarity: 50.2%
Query Match: 50.6%

Length: 1198
Matches: 155
Conservative: 58
Mismatch: 71
Indels: 25

```

DB:          9          Gaps:          4
US-10-634-548-2 (1-304) x US-11-096-568A-19558 (1-1198)
QY          18 GlyAenAanSerLeuThrThrHisArgPhe----- 27
DB          176 GGTGGAAATCCGAGATTCCTCGCTACCTCGGCACATGCTGCAGCGGCGGCTGG 235
QY          28 -----CysSerProGlyPheLeuLeu-----SerSerProCysPheLeuLeu 42
DB          236 ACCGCGCGCGCATCCCGCAACTCGCTGCTCTCGCGCTCTCCGCCACCGCGCGG 295
QY          43 ThrGlyMetGlySerAlaThrGlnLeuArgAlaArgSerLeuIleSer----- 59
DB          236 CTCGCGCGCTTCCCGGAAAGCTCCATCGG---AGCGCGCTCTCTCGCGCGCGCAC 352
QY          60 -----SerAlaValAlaThrAanSerLeuHisIleuSerGlyLeuVal 73
DB          353 CCGGCGGCTGCGCGGCTGCGCGCGCGCGCGCGCGCTGCGAGGACGAGCGGCTC 412
QY          74 ThrValAlaValLeuGlyValAlaValLeuValLeuSerPheGlySerLeuThrIys 93
DB          413 ACTGTGCTCATCACCGCGCGCGCTACTCTCTTGTCCGCGCTTGCAGACGACTCACCG 472
QY          94 ArgAenValIleGlnIleuSerLeuSerArgIleValIleIleuSerGlyLeuVal 113
DB          473 CCGGCGGCTCATGAAAGAGCTTGAAGCAGAAAGTTGTGACGTGATCCGCGCTCTG 532
QY          114 PheValIleuAlaTrpProIlePheSerGlySerThrGlnAlaArgIlePheAlaIle 133
DB          533 TTCATGTCACTTGGCCCTGTTACAGCAATTCGACAGAACGACGAGTATTCGCGCGCT 592
QY          134 ValProLeuValAenGlyLeuArgLeuValIleAenGlyLeuSerIleSerProAanSer 153
DB          593 GTCCCGCTTCTGAACTCCATGAGGCTTCTGATATATGACTCGTCTTACACTGATGA 652
QY          154 MetLeuIleLeuSerValThrArgGlnIleArgAlaGlnIleuLeuIysGlyProLeu 173
DB          653 GCTCTGTGTAATCACTGACACGTGAAGGAAACCGAGGAATGTGTGAGAGGTCCACTC 712
QY          174 PheTrpValIleuAlaLeuLeuPheSerAlaValPhePheTrpArgGlySerProIleGly 193
DB          713 TATATGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 772
QY          194 MetIleSerLeuAlaMetMetCysGlyGlyAenGlyIleAlaAanIleMetGlyArgIys 213
DB          773 ATCGTCTCTTGTGATGATGAGCGGCGGATGTTTGTGCTGACATTTGTGGAGAGG 832
QY          214 PheGlySerThrIysIleProIleProIleProIleProIleProIleProIleSerMet 233
DB          833 TATGCTTCAAGCAAGCTCCATTCATTCGAAAGAGAGAGCTGGCGCGGAGCANTCTGATG 892
QY          234 PheIlePheGlyPhePheIleSerIleAlaLeuLeuTrpTrpTrpSerSerLeuGlyTrp 253
DB          893 TTCATTTCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 952
QY          254 LeuHisMetAanTrpGlnTrpThrLeuGlnArgValAlaMetValSerMetValAlaTrp 273
DB          953 ATTGATGTTATCTGGAGAGGACCTTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1012
QY          274 ValValGlySerLeuProIleThrArgGlnLeuAanPheAanIleSerValProLeuAla 293
DB          1013 GTAGTGAAGGCTTCTCTGACTGAAAGTTGAGACAACTATCTGTTCTTTGGCC 1072
QY          294 ThrIleLeuAlaAlaTrpLeuSerPhe 302
DB          1073 ACCATGCTGGTAGCTTTCTCTGTTT 1099

```

```

; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 1769
; LENGTH: 986
; TYPE: DNA
; ORGANISM: Zea mays subsp. mays
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)..(986)
; OTHER INFORMATION: Cereb Seq. ID no. 15178973
US-11-096-568A-1769

Alignment Scores:
Pred. No.: 5,486-55 Length: 986
Score: 613.00 Matches: 125
Percent Similarity: 61.4% Conservative: 50
Best Local Similarity: 43.9% Mismatches: 54
Query Match: 39.9% Indels: 56
DB: Gaps: 4

US-10-634-548-2 (1-304) x US-11-096-568A-1769 (1-986)
QY          29 SerProGlyPheLeuLeu-----SerSerProCysPheIleGlyLeuThrGlyMetGly 46
DB          68 TCCCCCACTGCTGCTCTCTCTCGCGCTGCTGCCATCCCAAGCCGCGCTGCGGCTTGG 127
QY          47 SerAlaThrGlnLeuArgAlaArgSerLeuIleSer----- 59
DB          128 CCGGAGGACTCATGACGG---AGCGGCTCTGCTCGGCGTGGAGACCCCGCGCGGCGG 184
QY          60 -----SerAlaValAlaThrAanSerLeuHisIleuSerGlyValAlaThrValAlaVal 77
DB          185 GCGTGGCGCGCGCGCGCGCGCGCGCGCGCGCTGCGAGACGCGGCGCTCATCTGTCTATC 244
QY          78 LeuGlyGlyAlaTrpAlaLeuValLeuSerPheGlySerLeuThrIysArgAenValIle 97
DB          245 ACCGCGCGGCTCATCTCTTGTGCGCGCTTTCAGACAGCTTCCAGAGCGGCGGCTCATC 304
QY          98 GlnIleSerLeuSerArgIleLeuValHisIleuSerGlyLeuLeuPheValLeuAla 117
DB          305 GAAAGAGCTTGAAGAGAGAGTGTGACACGTCTATTCGGGCTCTGTTCAATGTCATCT 364
QY          118 TrpProIlePheSerGlySerThrGlnAlaArgIlePheAlaAlaPheValProLeuVal 137
DB          365 TGGCCCTGCTTC----- 376
QY          138 AenGlyLeuArgLeuValIleAenGlyLeuSerIleSerProAanSerMetLeuIleIys 157
DB          376 ----- 376
QY          158 SerValThrArgGlnIleArgAlaGlnIleuLeuIysGlyProLeuPheTrpValLeu 177
DB          377 -----AGGGAATGCTGAGAGGCTCCACTTATATATGTCCTG 412
QY          178 AlaLeuLeuPheSerAlaValPhePheTrpArgGlySerProIleGlyMetIleSerLeu 197
DB          413 GTGCTGCTGTGACGCTTGTAGCTTCTGCGGTGAGTCCCATCGGAAATCGTCTCTG 472
QY          198 AlaMetMetCysGlyGlyAenGlyIleAlaAanIleMetGlyArgIysPheGlySerThr 217
DB          473 TCGATGATAGCGGTGGCGCATGGTGTGCTGACATTTGTGGAGAGGATAGGCTACGCG 532
QY          218 IysIleProIleProIleProIleProIleProIleProIleProIleProIleSerMetPheIlePheGly 237
DB          533 AAGCTGCATTCATTCGAAAGAGCTGCGCGGAGCATTCGATTCATTTCTGCTGCT 592
QY          238 PhePheIleSerIleAlaLeuLeuTrpTrpTrpSerSerLeuGlyTrpLeuHisMetAan 257

```

```

Db      593 TTCTGCTGTCGCCGATGATGCTTCTCTCAAGCCGCGTTACATGATGTATTC 652
Qy      258 TTPGUTRThrLeuGlnArgValAlaMetValSerMetValAlaThrValAlaGluSer 277
Db      653 TGGGAAGAGCAGCTGTGTAAGCTGGCGCTGTGCACTGACAGCAGCACTAATGTGAGTGC 712
Qy      278 LeuProIleThrAspGlnLeuAspAlaSerValIleSerValProLeuAlaThrIleLeuAla 297
Db      713 GTTCTCTGACCCGAAAGCTGTATGACATATCTGTTCTTGGCCACCATGCTGGTA 772
Qy      298 AlaTyrLeuSerPhe 302
Db      773 GCTTTCTCTTGT 787

RESULT 3
US-11-096-568A-14811
; Sequence 14811, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 14811
; LENGTH: 1048
; TYPE: DNA
; ORGANISM: Zea mays subsp. mays
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(1048)
; OTHER INFORMATION: Ceres Seq. ID no. 1233653
US-11-096-568A-14811

Alignment Scores:
Pred. No.: 1,366-49 Length: 1048
Score: 562.50 Matches: 115
Percent Similarity: 60.4% Conservative: 53
Best Local Similarity: 41.4% Mismatches: 101
Query Match: 36.6% Indels: 9
DB: 9 Gaps: 4

US-10-634-548-2 (1-304) x US-11-096-568A-14811 (1-1048)
Qy      28 CysSerProGlyPheLeuIleSerSerProCysPheIleGlyLeuThrGlyMetGlySer 47
Db      92 TGTGGCGCT-----CGCTCAGCTCTCTCCGCCGCCGCGGACGAGTGTCTC 139
Qy      48 AlaThrGlnLeuArgAlaArgArgSerLeuIleSerSerAlaVal-----AlaThr 64
Db      140 GCAAGCGAG--CGAGCCGCGAGGCTCAACATGAGCGGCAATCTCGCTGAGGCCGCGC 196
Qy      65 AsnSerLeuLeuHisAspValGlyAlaThrValAlaValIleGlyValAlaTyrAlaLeu 84
Db      197 GGGGCGCTGGCCCACTGAGCTGAGTGTGCTGACAGTGTGCTGCTGCTGCTGCTGCTG 256
Qy      85 ValLeuSerPheGluSerLeuThrIlyArgAsnValIleGlnGlnSerLeuSerArgIys 104
Db      257 CTCAAATCTTTGAGAGAGCTCGCAAGCGCGGCTTCGACAGCAAACTCAGCAGGAAG 316
Qy      105 LeuValHisIleLeuSerGlyLeuLeuPheValIleAlaTyrProIlePheSerGlySer 124
Db      317 CTGTTCATATTAAGCTGGGCTGTGATTCATGCTTTTGGCTCTTTTAAGCTCTGGA 376
Qy      125 ThrGlnAlaArgTyrPheAlaAlaPheValProLeuValAsnGlyLeuArgLeuValIle 144
Db      377 TGGTATCTCTCTTCTCTGCTGCTGCTGACACAGGGGTTAATTAAGATGCTTCTTA 436
Qy      145 AsnGlyLeuSerIleSerProAsnSerMetLeuIleIleYserValThrArgGlnGlyArg 164
Db      437 CTGGGGCTGGGACTTATGAAAACGAGGCTATGTAAATCAATGAGCGGCTCTGGAAT 496

```

```

Qy      165 AlaGlnIleuLeuIysGlyProLeuPheTyrValIleAlaLeuPheSerAlaVal 184
Db      497 TACAGGAACTTCTCAAGGCCCTGCTATATGCTGCTACTATCACTTTGCAAGCTCT 556
Qy      185 PhePheTyrArgGluSerProIleGlyMetIleSerLeuAlaMetMetCysGlyIysAsp 204
Db      557 CTTTATGAGAACTCACTCACTGCTTATGACATTAATGCAACTTATGCTGCGAAT 616
Qy      205 GlyIleAlaAspIleMetGlyArgIysPheGlySerThrIlyIleProTyrAsnProArg 224
Db      617 GGTATAGCCGATGTGTGGAGAGCGCTAGGTAAGAAAGCTTCATACATCCCAAC 676
Qy      225 LysSerTPAlaGlySerIleSerMetIlePheGlyPhePheIleSerIleAlaLeu 244
Db      677 AATCATATGCTGGAAGCATGCAATGAGCGGCTGTCTTCTGCGCTCACTGAGGAT 736
Qy      245 LeuTyrTyrTyrSerSerLeuGlyTyrIleuHisMetAsnTyrGluThrThrLeuGlnArg 264
Db      737 ATGCATTACTTCCACACTTTCGTTTATGAGAAACATGATGATGCTTTAAGC--- 793
Qy      265 ValAlaMetValSerMetValAlaThrValAlaGluSerLeuProIleThrAspGlnLeu 284
Db      794 TTCTCGTGTCTCTGTAGCTGACACACTGCTGAATGCGACCCCATCAGCAGAACTT 853
Qy      285 AspAspAsnIleSerValProLeuAlaThrIleLeuAlaAlaTyrLeuSerPhe 302
Db      854 GATGCAATTTGACTGTCTTTGACATCATCTTCACTTGTGAGCTCATCTTT 907

RESULT 4
US-11-096-568A-14317
; Sequence 14317, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 14317
; LENGTH: 481
; TYPE: DNA
; ORGANISM: Zea mays subsp. mays
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(481)
; OTHER INFORMATION: Ceres Seq. ID no. 3626194
US-11-096-568A-14317

Alignment Scores:
Pred. No.: 9,796-41 Length: 481
Score: 475.00 Matches: 85
Percent Similarity: 83.8% Conservative: 29
Best Local Similarity: 62.5% Mismatches: 22
Query Match: 30.9% Indels: 0
DB: 9 Gaps: 0

US-10-634-548-2 (1-304) x US-11-096-568A-14317 (1-481)
Qy      167 GluLeuLeuYsGlyProLeuPheTyrValIleAlaLeuPheSerAlaValPhePhe 186
Db      26 GAATGCTGGAAGAGCTCACTATATATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 85
Qy      187 TTPArgGluSerProIleGlyMetIleSerLeuAlaMetMetCysGlyIysAspGlyIle 206
Db      86 TGGCGTATGCTCCCATGCGGATGCTCTCTTGTGATGATGAGCGGAGGATGTTT 145
Qy      207 AlaAspIleMetGlyArgIysPheGlySerThrIlyIleProTyrAsnProArgIysSer 226
Db      146 GCTGCAATGTGTGGAGAGGATATGCTCAAGCAAGCTGCATTCATCGAAGAAAGAGC 205

```

```

Qy 227 TrpAlaGlySerIleSerMetPheIlePheGlyPhePheIleSerIleAlaLeuLeuYr 246
Db 206 TGGGGCGGAGACATTCATTTCTTCTGTTCCGCTCCGCTATGATGATGCTC 265
Qy 247 TyrTyrSerSerLeuGlyTyrIleuHisMetAenTPGluThrThrLeuGlnArgValAla 266
Db 266 TACTCTCAAGCGCTGGGTACATGATGTATCTGGGAAGGACCTTGCTGTAAGCTGCG 325
Qy 267 MetValSerMetValAlaThrValAlaGluSerLeuProIleThrAspGlnLeuAspAsp 286
Db 326 CTGTGTGACACTGACGACACAGTACGTGAGTGGCTTCCTGTCGACCGAAGTTGATGATGAC 385
Qy 287 AsnIleSerValProLeuAlaThrIleLeuAlaAlaTyrLeuSerPhe 302
Db 386 AACATATCTGTTCTTGGCCACCATGCTGTAGCTTTCTTGTGTT 433

RESULT 5
US-10-932-182A-81375
; Sequence 81375, Application US/10932182A
; Publication No. US20060046253A1
; GENERAL INFORMATION:
; APPLICANT: NAKAO, YOSHIHIRO
; APPLICANT: NAKAMURA, NORIHISA
; APPLICANT: KODAMA, YUKIKO
; APPLICANT: FUJIMURA, TOMOKO
; APPLICANT: ASHIKARI, TOSHIHIKO
; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
; FILE REFERENCE: 030685-043
; CURRENT APPLICATION NUMBER: US/10/932,182A
; CURRENT FILING DATE: 2004-09-02
; NUMBER OF SEQ ID NOS: 197023
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 81375
; LENGTH: 978
; TYPE: DNA
; ORGANISM: Saccharomyces pastorianus
US-10-932-182A-81375

Alignment Scores:
Pred. No.: 4,23e-06 Length: 978
Score: 152.50 Matches: 59
Percent Similarity: 43.0% Conservative: 40
Best Local Similarity: 25.7% Mismatches: 84
Query Match: 9.9% Indels: 47
DB: 7 Gaps: 8

US-10-634-548-2 (1-304) x US-10-932-182A-81375 (1-978)
Qy 80 GYAlATyRAlAlLeuValLeuSerPheGluSerLeuThrLyAsnValIleGlnIn 99
Db 355 GCGCTTATTACTTTGTAAGCTGATACGCGGATATTAAATAAATGTTTATGGCCT 414
Qy 100 SerLeuSerArgLyLeuValHisIleLeuSerGlyLeuLeuPheValLeuAlaTrpPro 119
Db 415 TTGATATACGCGCTTCATCATATTGTTATTGATCTGATA---AGACTTAAACGCGCC 471
Qy 120 IlePheSerGlySerThrGlnAlaArgYrPheAlaAlaPheValProLeuValAsnGly 139
Db 472 TTTTTC----- 477
Qy 140 LeuArgLeuValIleAsnGlyLeuSerIleSerProAsnSerMetLeuIleYsSerVal 159
Db 478 -----AATATGCTTACTGTGAAGCTG 501
Qy 160 ThrArgGluGlyArgAlaGluGluLeu-----LeuLyGlyProLeuPheTyrValLeu 177
Db 502 GGTGGCGCTATATGAGAAAAGGAGATTTCATACATACATGAGGATATGTTGATACATCTT 561
Qy 178 AlaLeuLeuPheSerAlaValPhePheTyrPargGluSerProIleGlyMetIleSerLeu 197
Db 562 GGGTATATCTTTTCTTTTAACTTTTCTCTAAAGAT-----GTATACCTTAATATGTTA 615
Qy 198 AlaMetMetCysGlyGlyAspGlyIleAlaAspIleMetGlyArgLyAspPheGlySerThr 217

```

```

Db 616 TTTTGTCAAGTTGGTCCGATACAGCGCCCAACTATTTGAGAAAGTATGTCATTTA 675
Qy 218 LysIleProTyrAsnProArgLySerThrPlagIlySerIleSerMetPheIlePheGly 237
Db 676 ACACCCAAATTTGGCAGAAATAAATCCCTTCAGGTTTGATAGCTGCGCTTACACTGCT 735
Qy 238 PhePheIleSerIleAlaLeuLeuTyrTyrTyrSerSerLeuGlyTyrIleuHisMetAsn 257
Db 736 ---GTTATCACCTGCTGGTATTTTATGCGTATTTTGTTCCTGCTACAGCTAGCTAAC 792
Qy 258 -----TTP-----GluThrIleLeuGlnArgValAlaMetValSer 269
Db 793 AAACCTGGCAGATTCATGAGGCCAGAAACAGACATTTAAGTTGAATATGCTATCC 852
Qy 270 MetValAlaThrValAlaGluSerLeuProIleThrAspGlnLeu-----AspAsp 286
Db 853 TTGTTAGGTGGTGTGATGCTGCTTGTAGTGAAGGTATGATTTTCAACTGGGATGAT 912
Qy 287 AsnIleSerValProLeuAlaThrIleLeu 296
Db 913 AATTCACTATTCCTGCTGTCATCACTT 942

RESULT 6
US-10-932-182A-81375
; Sequence 81375, Application US/10932182A
; Publication No. US20060046253A1
; GENERAL INFORMATION:
; APPLICANT: NAKAO, YOSHIHIRO
; APPLICANT: NAKAMURA, NORIHISA
; APPLICANT: KODAMA, YUKIKO
; APPLICANT: FUJIMURA, TOMOKO
; APPLICANT: ASHIKARI, TOSHIHIKO
; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
; FILE REFERENCE: 030685-043
; CURRENT APPLICATION NUMBER: US/10/932,182A
; CURRENT FILING DATE: 2004-09-02
; NUMBER OF SEQ ID NOS: 197023
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 81375
; LENGTH: 978
; TYPE: DNA
; ORGANISM: Saccharomyces pastorianus
US-10-932-182A-81375

Alignment Scores:
Pred. No.: 4,23e-06 Length: 978
Score: 152.50 Matches: 59
Percent Similarity: 43.0% Conservative: 40
Best Local Similarity: 25.7% Mismatches: 84
Query Match: 9.9% Indels: 47
DB: 7 Gaps: 8

US-10-634-548-2 (1-304) x US-10-932-182A-81375 (1-978)
Qy 80 GYAlATyRAlAlLeuValLeuSerPheGluSerLeuThrLyAsnValIleGlnIn 99
Db 355 GCGCTTATTACTTTTACTTGTATACGCGGATATTAAATAAATGTTTATGGCCT 414
Qy 100 SerLeuSerArgLyLeuValHisIleLeuSerGlyLeuLeuPheValLeuAlaTrpPro 119
Db 415 TTGATATACGCGCTTCATCATATTGTTATTGATCTGATA---AGACTTAAACGCGCC 471
Qy 120 IlePheSerGlySerThrGlnAlaArgYrPheAlaAlaPheValProLeuValAsnGly 139
Db 472 TTTTTC----- 477
Qy 140 LeuArgLeuValIleAsnGlyLeuSerIleSerProAsnSerMetLeuIleYsSerVal 159
Db 478 -----AATATGCTTACTGTGAAGCTG 501
Qy 160 ThrArgGluGlyArgAlaGluGluLeu-----LeuLyGlyProLeuPheTyrValLeu 177

```


Db 502 GGTGGCTAATGAGAAAAAGAGATTACATACATCAATGGGTAATTGTGTACATTTT 561
 Qy 178 AAlaLeuPheSerAlaValaPhePheTpaArgGluSerProIleGlyMetIleSerLeu 197
 Db 552 GGGTAACTCTTTCTTTAACTTTCTCTTAAAGAT-----GTTACCTTAATATCGTTA 615
 Qy 198 AlameMetCyGlyIleAapGlyIleAlaAapIleMetGlyArgIlePheGlySerThr 217
 Db 616 TTTTGTCTAAGTGTGTCCGATACAGCCCGCAACTATTGAAAGAAAGATGTCTATTAA 675
 Qy 218 LysIleProTyraProArgIleSerThrAlaGlySerIleSerMetPheIlePheGly 237
 Db 676 AACACCAATTGGCAAGAAATTAATCCCTTCAGAGTTGATAGTCGCGTTTACAGTTGT 735
 Qy 238 PhePheIleSerIleAlaLeuLeuIleTyTyTyTySerSerLeuGlyTyIleuHisMetLeu 257
 Db 736 ---GTTATCACTGCTGTGGTATTATTATGCTATTGTTCTCGCTACAGCTACGTCAC 792
 Qy 258 -----Trp-----GluThrThrLeuGlnArgValaAlaMetValSer 269
 Db 793 AAACCTGGCAGATCCATCGAGCCCAAGAAACAGACAGTTAAGTTGAAATATGCTATCC 852
 Qy 270 MetValaAlaThrValaGluSerLeuProIleThrAspGlnLeu-----AspAsp 286
 Db 853 TTGTAGGTGGTGTGTAGTGTCTTTGAGTGAAGGATAGATTGTTCAACTGGGATGAT 912
 Qy 287 AsnIleSerValProLeuAlaThrIleLeu 296
 Db 913 AATTCACTATTCTCTCTCTGTCATCACTT 942

RESULT 7

US-10-932-182A-174337
 ; Sequence 174337, Application US/10932182A
 ; Publication No. US20060046253A1
 ; GENERAL INFORMATION:
 ; APPLICANT: NAKAO, YOSHIHIRO
 ; APPLICANT: NAKAMURA, NORIHISA
 ; APPLICANT: KODAMA, YUKIKO
 ; APPLICANT: FUJIMURA, TOMOKO
 ; APPLICANT: ASHIKARI, TOSHIHIKO
 ; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
 ; FILE REFERENCE: 030685-043
 ; CURRENT APPLICATION NUMBER: US/10/932,182A
 ; CURRENT FILING DATE: 2004-09-02
 ; NUMBER OF SEQ ID NOS: 197023
 ; SOFTWARE: PatentIn version 3.3
 ; SEQ ID NO 174337
 ; LENGTH: 525
 ; TYPE: DNA
 ; ORGANISM: Saccharomyces pastorianus
 US-10-932-182A-174337

Alignment Scores:

Pred. No.: 3,51e-06 Length: 525
 Score: 150.00 Matches: 52
 Percent Similarity: 42.64 Conservative: 31
 Best Local Similarity: 26.74 Mismatches: 64
 Query Match: 9.8% Indels: 48
 DB: 7 Gaps: 6

US-10-634-548-2 (1-304) x US-10-932-182A-174337 (1-525)

Qy 116 LeuAlaTppProIlePheSerGlySerThrGlnAlaArgTyrrPheAlaAlaPheValPro 135
 Db 10 TTAACCTGGCCCTTTTTC----- 27
 Qy 136 LeuValaGnglyLeuArgLeuValIleAanglyLeuSerIleSerProAsnSerMetLeu 155
 Db 28 -----AATATGCTTAT 39
 Qy 156 IleYsSerValThrArgGnglyArgAlaGnglyLeu-----LeuYsGlyProLeu 173
 Db 40 TGTAGCAGGTAGGTGCTTATATGAGAAAAAGAGATTACACGTACACGCGGTATTG 99

Qy 174 PheTyValIleAlaLeuLeuPheSerAlaValaPhePheTpaArgGluSerProIleGly 193
 Db 100 TGTAACTCTCTGGGTGTGTGTTCTTTAGTTTCTTTCCAAAGAC-----GTTACT 153
 Qy 194 MetIleSerLeuAlaMetMetCyGlyIleAapGlyIleAlaAapIleMetGlyArgIle 213
 Db 154 TTGATATATCTTCTTCTACTAGATTGTCCGATACAGCCCGCAACATCGAAGAA 213
 Qy 214 PheGlySerThrLysIleProTyraProArgIleSerThrAlaGlySerIleSerMet 233
 Db 214 TATGTCTATTAAACACCAATTGGCGAAGAAATCTCTTCAGGCTCGATACCTGCA 273
 Qy 234 Phe-----IlePheGlyPhePheIleSerIleAlaLeu 244
 Db 274 TTACAGTTGGGGCATCATCCTGTTCACATTAAATTTATTTT-----ATACCACT 327
 Qy 245 LeuTyTyTyTySerSerLeuGlyTyIleuHisMetLeuThrThrLeuGlnArg 264
 Db 328 TACTACTACGTCAATTAACCTGAGAGATTGAATGACTCCGAAACAGTAATTAAC 387
 Qy 265 ValaAlaMetValSerMetValaAlaThrValaGluSerLeuProIleThrAspGlnLeu 284
 Db 388 TTGAATATGTTATCGTGTAGTGTGTGCTGCTGCATTTGACGGAAGTATTGATTG 447
 Qy 285 -----AspAsnIleSerValProLeuAlaThrIleLeu 296
 Db 448 TTCACTGGAGTACACTTCACTATTCTCTGATCATCACTT 492

RESULT 8

US-10-932-182A-174337
 ; Sequence 174337, Application US/10932182A
 ; Publication No. US20060046253A1
 ; GENERAL INFORMATION:
 ; APPLICANT: NAKAO, YOSHIHIRO
 ; APPLICANT: NAKAMURA, NORIHISA
 ; APPLICANT: KODAMA, YUKIKO
 ; APPLICANT: FUJIMURA, TOMOKO
 ; APPLICANT: ASHIKARI, TOSHIHIKO
 ; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
 ; FILE REFERENCE: 030685-043
 ; CURRENT APPLICATION NUMBER: US/10/932,182A
 ; CURRENT FILING DATE: 2004-09-02
 ; NUMBER OF SEQ ID NOS: 197023
 ; SOFTWARE: PatentIn version 3.3
 ; SEQ ID NO 174337
 ; LENGTH: 525
 ; TYPE: DNA
 ; ORGANISM: Saccharomyces pastorianus
 US-10-932-182A-174337

Alignment Scores:

Pred. No.: 3,51e-06 Length: 525
 Score: 150.00 Matches: 52
 Percent Similarity: 42.64 Conservative: 31
 Best Local Similarity: 26.74 Mismatches: 64
 Query Match: 9.8% Indels: 48
 DB: 7 Gaps: 6

US-10-634-548-2 (1-304) x US-10-932-182A-174337 (1-525)

Qy 116 LeuAlaTppProIlePheSerGlySerThrGlnAlaArgTyrrPheAlaAlaPheValPro 135
 Db 10 TTAACCTGGCCCTTTTTC----- 27
 Qy 136 LeuValaGnglyLeuArgLeuValIleAanglyLeuSerIleSerProAsnSerMetLeu 155
 Db 28 -----AATATGCTTAT 39
 Qy 156 IleYsSerValThrArgGnglyArgAlaGnglyLeu-----LeuYsGlyProLeu 173
 Db 40 TGTAGCAGGTAGGTGCTTATATGAGAAAAAGAGATTACACGTACACGCGGTATTG 99


```

Db      229 CGAATAATGGCATTTTC-----ATCATTTTCTTACTATATAACCGTCATTCGA 279
Qy      123 GlySerThrGluAlaArgTyrPheAlaAlaPheValProLeuValAsnGlyLeuArgLeu 142
Db      280 ATGCATTTCA-----AACCTTGCGTGAATAAT 303
Qy      143 ValIleAsnGly-----LeuSerIleSerProAsnSerMetIleuIleuSerVal 159
Db      304 GCATTCCTCCGACAAATACACAGCTCTTCTTATCAATTGAGTACATAGATTCAAAACCTTA 363
Qy      160 ThrArgGluGlyArgAlaGluGluLeu----- 169
Db      364 CCGCGCTTGAAGCTGATTGATTAACAATAAGAGTTGCTGATGATAGGAGCAC 423
Qy      170 LysGlyProLeuPheTyr---ValLeuAlaLeuLeuPhe-----SerIleValPhePhe 186
Db      424 AGCGGGCCATGATCATATCATATCATCTTATTTATCTTCCGAAATCAACACCTTTATTA 483
Qy      187 TrpArgGluSerProIleGlyMetIleSerLeuAlaMetMetCysGlyValAspGlyIle 206
Db      484 ATGAATACTCTCCAAATGGCTCTATATAGATTGGGAAT-----GGTATTCCTTA 534
Qy      207 AlaAspIleMetGlyArgGlyPheGlySerThrLysIleProTyrAsnProArgLysSer 226
Db      535 GCATCTATATATGTAAGGATAGCGCATTCGTTGGAAA---GTAACCAAAAAACT 591
Qy      227 TrpAlaGlySerIleSerMetPheIlePheGlyPhePheIleSerIleAlaLeuLeuTyr 246
Db      592 TTAGAGGGAACCTCTGCTTTATATAGTAAAGATTTCGTTTGTATATATGTATAGT 651
Qy      247 TyrTyrSerSerLeuGlyTyrLeuHisMetAsnTrpGluThrThrLeuGluArgValAla 266
Db      652 TTGTATTAAGCTGCATTTTAAACCACTTA-----ACTACTTTCAAAATGGTTACC 702
Qy      267 MetValSerMetValAlaThrValValGluSerLeuProIleThrAspGlnLeuAspAsp 286
Db      703 CTTGTATACACTG---AGTGAGAGTGTAGAGGTAATAGTGTG-----CTTATATGAC 750
Qy      287 AsnIleSerValPro 291
Db      751 AATATTTTGAATACCT 765

RESULT 11
US-10-932-182A-1952
; Sequence 1952, Application US/10932182A
; Publication No. US20060046253A1
; GENERAL INFORMATION:
; APPLICANT: NAKAO, YOSHIHIRO
; APPLICANT: NAKAMURA, NORIHISA
; APPLICANT: KODAMA, YUKIKO
; APPLICANT: FUJIMURA, TOMOKO
; APPLICANT: ASHIKARI, TOSHIHIKO
; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
; FILE REFERENCE: 030685-043
; CURRENT APPLICATION NUMBER: US/10/932,182A
; CURRENT FILING DATE: 2004-09-02
; NUMBER OF SEQ ID NOS: 197023
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 1952
; LENGTH: 1560
; TYPE: DNA
; ORGANISM: Saccharomyces pastorianus
US-10-932-182A-1952

Alignment Scores:
Pred. No.: 6,95e-05 Length: 1560
Score: 143.50 Matches: 56
Percent Similarity: 46.2% Conservative: 48
Best Local Similarity: 24.9% Mismatches: 78
Query Match: 9.3% Indels: 43
DB: 7 Gaps: 9
US-10-634-548-2 (1-304) x US-10-932-182A-1952 (1-1560)

```

```

Qy      84 LeuValIleSerPheGluSer---LeuThrLysArgAsnValIleGlnGlnSerLeuSer 102
Db      925 CTACGCTATCCCGTACCAAGATATTGTATGCAAAAGACAGTCAATCGTTAAATCCCTCC 984
Qy      103 ArgLysLeuValHisIleLeuSerGlyLeuLeuPheValIleuAlaTrpProIlePheSer 122
Db      985 CGAAGAGTGTGGCACTTCAATTAATATCTTATCTTATAGTCCATCATCTTCCAAATAGATTCA 1044
Qy      123 GlySerThrGluAlaArgTyrPheAlaAlaPheValProLeuValAsnGlyLeuArgLeu 142
Db      1045 -----AATTTGTGAATAAT 1059
Qy      143 ValIleAsnGlyLeu-----SerIle 149
Db      1060 GCATTCCTGGAACAAATACCAAGCTCTTCTTATCATGTTGAATATATAGTTTCAAAACCTTA 1119
Qy      150 SerProAsnSerMetLeuIleLysSerValThrArgGluGlyArgAlaGluGluLeu 169
Db      1120 CCGCTCTAGAGATCTCTATATGAACTACAACTTAGAAGCTTTCGCGAGCCAGACCAT 1179
Qy      170 LysGlyProLeuPheTyr---ValLeuAlaLeuLeuPhe-----SerIleValPhePhe 186
Db      1180 AGTGGGCCCTGATCATATCTTATCTTATATATATTTGGAATATCGATACCTTGTG 1239
Qy      187 TrpArgGluSerProIleGlyMetIleSerLeuAlaMetMetCysGlyValAspGlyIle 206
Db      1240 GTGAATGACTGCCCATGGAATTAATAGTTGGGAAT-----GGGCAATTCCTGTG 1290
Qy      207 AlaAspIleMetGlyArgGlyPheGlySerThrLysIleProTyrAsnProArgLysSer 226
Db      1291 GCATCTATATATCGTAAAGATATGTCATAATTCGTTG---CGAGTACCCAAAGACT 1347
Qy      227 TrpAlaGlySerIleSerMetPheIlePheGlyPhePheIleSerIleAlaLeuLeuTyr 246
Db      1348 ATAGAGGGAACCTCTGCGTCAACATTAAGATTATTTGTTGTTGTTGTTGTTACTTAC 1407
Qy      247 TyrTyrSerSerLeuGlyTyrLeuHisMetAsnTrpGluThrThrLeuGluArgValAla 266
Db      1408 TTGTATTAAGCAACCGTGTTCATGCAATTTA-----ACCGTTTCAAAATGGCTT 1455
Qy      267 MetValSerMetValAlaThrValValGluSerLeuProIleThrAspGlnLeuAspAsp 286
Db      1456 TGTGTTGTATATTAACCGGGGTATTTGGAAGAAATAGTGTG-----CTAATATGAT 1506
Qy      287 AsnIleSerValPro 291
Db      1507 AATATTTTGAATACCT 1521

RESULT 12
US-10-932-182A-1952
; Sequence 1952, Application US/10932182A
; Publication No. US20060046253A1
; GENERAL INFORMATION:
; APPLICANT: NAKAO, YOSHIHIRO
; APPLICANT: NAKAMURA, NORIHISA
; APPLICANT: KODAMA, YUKIKO
; APPLICANT: FUJIMURA, TOMOKO
; APPLICANT: ASHIKARI, TOSHIHIKO
; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
; FILE REFERENCE: 030685-043
; CURRENT APPLICATION NUMBER: US/10/932,182A
; CURRENT FILING DATE: 2004-09-02
; NUMBER OF SEQ ID NOS: 197023
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 1952
; LENGTH: 1560
; TYPE: DNA
; ORGANISM: Saccharomyces pastorianus
US-10-932-182A-1952

Alignment Scores:
Pred. No.: 6,95e-05 Length: 1560

```

Score: 143.50 Matches: 56
 Percent Similarity: 46.2% Conservative: 48
 Best Local Similarity: 24.9% Mismatches: 78
 Query Match: 9.3% Indels: 43
 Gaps: 9

US-10-634-548-2 (1-304) x US-10-932-182A-1952 (1-1560)

```

QY      84 LeuValLeuSerPheGluSer----LeuThrIysAspGlnValIleGlnGlnSerLeuSer 102
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      925 CTACGCTCATCCGTAACCAAGATATTTTCGAAAAGACGCTATCGTTAAATCCCTCC 984
QY      103 ArgIysLeuValHisIleLeuSerGlyLeuLeuPheValIleuValIleuValIleuVal 122
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      985 CGAAAAGTGGCGACTTCATATATCTTACTTATAGTCCATCATCTTCAATATAGATTC 1044
QY      123 GlySerThrGluAlaArgTyPheAlaIlePheValProLeuValIleuValIleuVal 142
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      1045 -----AATTTGTGMAAAAT 1059
QY      143 ValIleAsnGlyLeu-----SerIle 149
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      1060 GCATGTCTGGAAACATACCACTCTTATCATGTAATATATAGTTTCAAAACCTA 1119
QY      150 SerProAsnSerMetLeuIleLeuSerValThrArgGluGlyArgAlaGluGluLeu 169
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      1120 CCGCCTCATAGACTCTTATATGACTACATTAAGAGTTGCCGACCGACAGACCAT 1179
QY      170 LysGlyProLeuPheTyr---ValIleuAlaLeuLeuPhe---SerAlaValPhePhe 186
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      1180 AGTGGGCCCTCATATATCCATCTTATATATATTTGGAATATCATACCTTGTG 1239
QY      187 TrpArgGluSerProIleGlyMetIleSerLeuAlaMetMetCysGlyGlyAspGlyIle 206
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      1240 GTGAATGACTCCGCCATGGGATTAATAGGTTGGGAAT-----GGCATTCCTCG 1290
QY      207 AlaAspIleMetGlyArgIysPheGlySerThrIysIleProTyAsnProArgIysSer 226
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      1291 GCATCTATATATCGTAAAGATATATGTCATAATTCCTGG---CGAGTATCCAAAGACT 1347
QY      227 TrpAlaGlySerIleSerMetPheIlePheGlyPhePheIleSerIleAlaLeuLeuTy 246
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      1348 ATAGAGGGAACCTTGGCGTTACCATTAACAGTTTATGTTGGTATTAATCTTAC 1407
QY      247 TyrTyPheSerLeuGlyTyIleuHisMetLeuPheGluThrThleuGlnArgValAla 266
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      1408 TTTGATTAAGCAACCGTGTTCAGTCATTTA-----ACCGTTTACAAATGCTT 1455
QY      267 MetValSerMetValAlaThrValValGluSerLeuProIleThrAspGlnLeuAsp 286
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      1456 TGTGTTTGTACATTAAGCGGGTATTTGGAAGAAATAGTGTG-----CTAATATGAT 1506
QY      287 AsnIleSerValPro 291
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      1507 AATATTTGTATACCT 1521

RESULT 13
US-10-499-290-1
/ Sequence 1, Application US/10499290
/ Publication No. US20060051362A1
GENERAL INFORMATION:
/ APPLICANT: Jacques, Mario
/ APPLICANT: Coulton, James
/ APPLICANT: Mikael, Leonie G.
/ APPLICANT: Khameshan, Ali
/ APPLICANT: Siros, Marc
/ APPLICANT: Universite de Montreal
/ APPLICANT: McGill University
/ APPLICANT: Universite du Quebec a Trois-Rivieres
/ TITLE OF INVENTION: Antigens for Actinobacillus
/ FILE REFERENCE: 10662-113ep
/ CURRENT APPLICATION NUMBER: US/10/499,290

```

```

/ CURRENT FILING DATE: 2004-06-18
/ PRIOR APPLICATION NUMBER: US 60/340,483
/ PRIOR FILING DATE: 2001-12-18
/ NUMBER OF SEQ ID NOS: 41
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 1
/ LENGTH: 7974
/ TYPE: DNA
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: flu operon in Actinobacillus pleuropneumoniae
/ NAME/KEY: misc feature
/ LOCATION: (1)...(7974)
/ OTHER INFORMATION: n = A,T,C or G
US-10-499-290-1

Alignment Scores:
Pred. No.: 0.53 Length: 7974
Score: 115.50 Matches: 84
Percent Similarity: 39.5% Conservative: 53
Best Local Similarity: 24.2% Mismatches: 111
Query Match: 7.5% Indels: 99
Gaps: 19

US-10-634-548-2 (1-304) x US-10-499-290-1 (1-7974)
QY      40 IleGlyLeuThrGlyMetGlySerAlaThr-----GlnLeuArgAlaArgArg--- 55
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      3350 TTTGGTTTATTCGCTTGGCTTGGCAACTATTTGTTGCCCACTGCATATTCGCCGTTT 3409
QY      56 -----SerLeu 57
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      3410 AGTCATCACTATGATACGCTTTTGTACTGGGGCGGTATTTGGCGAATACGGAATTTG 3469
QY      58 IleSerSerAlaValAlaThrAsnSerLeuLeuHis---AspValGlyAlaThrValAla 76
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      3470 CTGTTCACGTTATTTGCTTTTATGCGGAGATTCATTTACCGACCGGTGGTGTGCA 3529
QY      77 ValLeuGlyGlyAlaTyAlaLeuValLeuSerPheGluSerLeuThrIysArgAsnVal 96
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      3530 TTATTAAGTACCGCGCTTGTCTTGGCTGATGTTAGTATTCGATTCGCCGATACCGGACGT 3589
QY      97 IleGlnGlnSerLeuSerArgIys-----LeuValHisIle 108
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      3590 TTACGCAAAAGGATGTCGCAAAAATTCGTCATTTGACACCGCTTGTCTTATTTGATA 1649
QY      109 LeuSerGlyLeuLeuPheValLeuAlaTrpProIlePheSerGly----- 123
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      3650 CTCGGCGGTTTACCTTGACAAATTTGCTTAGCGTATTTGTCGCAAAAGCATCAACGCG 3709
QY      124 -----SerThrGluAlaArgTyPheAlaIlePheValProLeuValIleuVal 140
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      3710 TGGCAATGACGCAAAAGACGATTAATTCATTCGCGCTTGCCT----- 3754
QY      141 ArgLeuValIleAsn-----GlyLeuSerIleSerProAsnSerMetLeuIleIys 157
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      3755 CCTTAATATGTAAGCAATAGCGGTGGAATTTACTGTGGTCCGCGTGTGATTTACAG 3814
QY      158 SerValThrArgGluGlyArgAlaGlu---GluLeuLeu-----LysGlyProLeu 173
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      3815 CGTTGACTTTAAACCCGATGGCAAGCCCGCAATATTTGGGGGTGTTCCGCGCATCA 3874
QY      174 PheTyPheAlaLeuAlaLeuLeuPhe-----SerAlaValPhePheTyPArg 188
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      3875 ATGGGGGTTTAAACCGCTTGTCTGTTGTGTGTCACAACAAACGAATGTTTGTG 3934
QY      189 GluSerProIle----- 192
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      3935 GAGAGGATTCCTTGGTCTTGTGTCGATTAATTTGATTAACGATCAATACGCGCAGC 3994
QY      193 GlyMetIleSerLeuAlaMetMetCysGlyGlyAspGlyIleAla-----AspIle 209
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

```

Db 3995 GGTATGTTCCGGAAAAAGTATTCGACGTGATTCGCTTATCGCATTTATTTGACT 4054
Qy 210 MetGlyArgLysPheGlySerThrIlePro---TyrAspProArgLysSerTrp--- 227
Db 4055 CTGCACGCTTTAGCATTGCGACGTGAGACCTTCGACGCCAATCAGCTGATTACGGAGCA 4114
Qy 228 AlaGlySerIle-----SerMetPheIlePheGlyPhePheIleSerIle 242
Db 4115 TCCGATCATCAGCAAAATGTTACTATGAAATGGCTATCGCTTCGCTACGGCTACGTTT 4174
Qy 243 AlaIleuLeuTyrTyrTyrSerSerLeuGlyTyrIleuHisMetAsnTrpGluThrIleu 262
Db 4175 GCACACTACTT-----TTTGCAGATTGG-----ATTTCAGCCGCTTG----- 4210
Qy 263 GlnArgValAlaMetValSerMetValAlaThrValAlaGluSerLeuProIleThrAsp 282
Db 4211 -----CTCGCCTTATCGCTTACAGCGCCGAGGACCAATCCCTGGTTAAATTTA 4264
Qy 283 Gln-----LeuAspAspAsnIleSerValProIleuAlaThrIle 295
Db 4265 AAACAGCTTCGCTGAGTTTATTTATTTAGTGGCTATTACGCGCATACGCACTTGG 4324
Qy 296 LeuAlaAlaTyrLeuSerPhe 302
Db 4325 GTTATCGCTCCGTTAAGTTT 4345

RESULT 14
US-10-793-626-2167
; Sequence 2167, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMBERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: P03480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: Patent Ver. 2.1
; SEQ ID NO 2167
; LENGTH: 1152
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: nucleic acid sequence
US-10-793-626-2167

Alignment Scores:
Pred. No.: 2.82 Length: 1152
Score: 98.50 Matches: 70
Percent Similarity: 40.94 Conservative: 58
Best Local Similarity: 22.44 Mismatches: 115
Query Match: 6.44 Indels: 71
Gaps: 13

US-10-634-548-2 (1-304) x US-10-793-626-2167 (1-1152)
Qy 13 GlnLeuCysArgPheGlyAsnAsnSerLeuThrThrIleAspGlySerProGlyPhe 32
Db 181 CAATCATGCTTTTAAATCAAGTACGCTGACAAATATTCATTTGAAACA----- 231
Qy 33 LeuIleSerSerProCysPheIleGlyLeuThrGlyMetGlySerAlaThrGlnLeuArg 52
Db 232 ATCATTTACTATCC-----TGGTTGACTGTTTATGAGCAAGCATATTTATTTGG 282
Qy 53 AlaArgArgSerLeuIleSerSerAlaValAlaThrAsnSerLeuLeuHisAspValGly 72
Db 283 ATTAACAATGTAATTTGTTATGTAGCTTTGAAAAAGGT-----TTGAAAAAGCTTCT 336
Qy 73 AlaThrValAlaValLeuGlyGlyAlaTyrAlaLeuValLeuSerPheGluSerLeuThr 92

```

```

Db 337 AAAATATGATGACCTCTATTATTATCTTTAAATTATCGTTGACAAATCTTAACT 396
Qy 93 LysArgAsnValIleGln-----GlnSerLeuSer 102
Db 397 TTAGAGAGTCTTTTAGAAGGTGTCATTATATCTGCAACCTCGAGTTGAAGATATGCT 456
Qy 103 ArgLysLeuValHisIleLeuSerGlyLeuLeuPheValLeuAlaTrpProIlePheSer 122
Db 457 -----ATTCAAGTCTACTATTGTGGCTTAGGACAATCGTTTTCAG 498
Qy 123 -----GlySerThrGluAlaArgTyrPheAlaIlePheValPro----- 135
Db 499 CTGTCCCTAAGTACACACCGAATGATGATCTATGCAACCTATGCACTAAAAATATGACG 558
Qy 136 -----LeuValAsnGlyLeuAlaGlyLeuValIleAsnGly 146
Db 559 ATAAAGTCTTCAGCATTTCAATTGTCGTAATGAAATATTTAAATTTCTGTCTGTGCTGCA 618
Qy 147 LeuSerIleSerProAsnSerMetLeuIleLysSerValThrArgGluGlyArgAlaGlu 166
Db 619 TTAGCTATATTTCTCGCGCTTAAAAACATTTGGTTACCAACCCCAAGAGGC----- 669
Qy 167 GluLeuLeuGlyGlyProLeuPheTyrValLeuAlaLeuLeuPheSerAlaValPhePhe 186
Db 670 -----CTGGCTTATTTATTTAAGGTTTACCACTAGATTTAAGCAAAATGAC-TTT 719
Qy 187 TTPArgGluSerProIleGlyMetIleSerLeuAlaMetCysGlyGlyAsp-GlyIle 206
Db 720 TGGTACATCTTTTACCTTTATATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 779
Qy 206 AlaAspIleMetGlyArgLysPheGlySer---ThrLysIleProTyrAsnProArgLys 225
Db 780 TATATCATTTATTAAGTAAATGATGATCTAATTTTACTAAATGATATAGTAAAGACA 839
Qy 225 sSerTrpAla-----GlySerIleSerMetPheIlePheGlyPhePheIleSerIleAl 243
Db 840 AAAGTGCATTCATGAGTAGTATGATATCTGATTTATC-----ATTAGATATCCC 887
Qy 243 AlaLeuLeuTyrTyrIleSerSerIleGlyTyrIleuHisMetAsnTrpGluThrIleuG 263
Db 888 AGCAACATTATCTTTGATGATCTAAGTCATTCGATTCGCTTTGGC----- 930
Qy 263 nArgValAlaMetValSerMetValAlaThrValAlaGluSerLeuProIleThrAspG 283
Db 931 -----GCTGGTACATATTTGATATATG-----GATTT 959
Qy 283 nLeuAspAspAsnIleSerValProLeuAlaThrIle 295
Db 960 TATGTATCTAATATTTCTTATGACCATTAAGGGGCACTA 996

RESULT 15
US-10-793-626-3427/c
; Sequence 3427, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMBERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: P03480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: Patent Ver. 2.1
; SEQ ID NO 3427
; LENGTH: 3108
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: nucleic acid sequence
US-10-793-626-3427

```

Alignment Scores:

Pred. No.:	10.1	Length:	3108
Score:	98.50	Matches:	70
Percent Similarity:	40.9%	Conservative:	58
Best Local Similarity:	22.4%	Mismatches:	115
Query Match:	6.4%	Indels:	71
DB:	8	Gaps:	13

US-10-634-548-2 (1-304) x US-10-793-626-3427 (1-3108)

Search completed: March 16, 2006, 23:57:14
 Job time : 440 secs

```

QY      13  GlnLeuCysArgPheGlyAsnAsnSerLeuThrHisArgPheCysSerProGlyPhe 32
DB      3087  CAATCATGCTTTTAAATCAAGTACGCTGACAAATATTCAATTGAAACA----- 3037
QY      33  LeuIleSerSerProCysPheIleGlyLeuThrGlyMetGlySerAlaThrGlnLeuArg 52
DB      3036  ATCATTTAGTAATCCA-----TGGTTGACTGTTTATAGTCGCAAGCATATTATTG 2986
QY      53  AlaArgSerSerIleSerSerAlaValAlaThrAsnSerLeuLeuHisArgValGly 72
DB      2985  ATACAAATGTAATTGTTATGTTAGTGTGAAAAAGCT-----TTAGAAAAAGCTTCT 2932
QY      73  AlaThrValAlaValLeuGlyValAlaThrAlaLeuValLeuSerPheGlySerLeuThr 92
DB      2931  AAAATATATGATGCTCTATTATTATTCTTTAATTATGCTGTAGACAAATCTTTAACT 2872
QY      93  LysArgSerValIleGln-----GlnSerLeuSer 102
DB      2871  TTAGAAAGTCTTTAGAAAGTGTACGTATATATACGCAACTCGAGTTGAAGATATGCT 2812
QY      103  ArgLysLeuValHisIleLeuSerGlyLeuLeuPheValLeuAlaTrpProIlePheSer 122
DB      2811  -----ATCAAGGTGACTATTGCTTAGGACAAATCGTTTTTACG 2770
QY      123  -----GlySerThrGlnAlaArgTyrPheAlaAlaPheValPro----- 135
DB      2769  CTGTCCCTAGTACAAACCGAATGATTTACTATGCAAGCTATGACCTTAAATATATGACG 2710
QY      136  -----LeuValaGnglyLeuArgLeuValIleGngly 146
DB      2709  ATAAAGCTTCAGACTTTCATTTGCGTAATGAATATTTAATTTCTGTGCTGCTGGA 2650
QY      147  LeuSerIleSerProAsnSerMetLeuIleLysSerValThrArgGlnGlyArgAlaGln 166
DB      2649  TTAGCTATTTCTCGCGCTTAAACATTTGGTTACCAACCCCAAGAAAGC----- 2599
QY      167  GlnLeuLeuLysGlyProLeuPheTyrValLeuAlaLeuLeuPheSerAlaValPhePhe 186
DB      2598  -----CTGGCTTATTATTAAAGTTTACCACTAGTATTGCGCAAAATGAC-TTT 2549
QY      187  TrpArgGlnSerProIleGlyMetIleSerLeuAlaMetMetCysGlyGlyAsp-GlyTl 206
DB      2548  TGGTACATTTCTTTACTTTATTTTACTATTATTCTTATTGCGGCATTAACGCTTTC 2489
QY      206  eAlaAspIleMetGlyArgLysPheGlySer--ThrLysIleProTyrAsnProArgLys 225
DB      2488  TATATCATTTATTAAGATTAAATATATATTTACTTAAATAATGATAATGATAAAGACA 2429
QY      225  sSerTrpAla-----GlySerIleSerMetPheIlePheGlyPhePheIleSerIleAl 243
DB      2428  AAAAGTGCAGCAATCATAGTATGTTACTGTGATTTATTC-----ATTAGTATCCC 2381
QY      243  AlaLeuLeuTyrTyrSerSerLeuGlyTyrLeuHisMetAsnTrpGlnThrThrLeuGln 263
DB      2380  AGCAACATTATCTTTTATAGTCTAAGTCATTTGCGCTTTGGC----- 2338
QY      263  nArgValAlaMetValSerMetValAlaThrValValGlnSerLeuProIleThrAspGln 283
DB      2337  -----GCTGTACGATATTGATAATATG-----GATTT 2309
QY      283  nLeuAspAspAsnIleSerValProLeuAlaThrIle 295
DB      2308  TATTGTATCTAATATTTCTTATGCAATTAGGGGACACTA 2272

```